Security Assurance of Disaster Recovery Data in Information Collection System Based on Artificial Intelligence Encryption Algorithm

Siyuan Suo*, Meiling Zhang, Jun Zhang and Jiayi Liu

Measurement Management Room, State Grid Marketing Service Center of Shanxi Province, Taiyuan 044399, Shanxi, China
E-mail: siyuan_suo15@outlook.com; meiling_zhang45@163.com; 1355387236@qq.com; jun_zhang98@123.com; 343738836@qq.com; jiayi_liu90@163.com; ljy4631@126.com
*Corresponding Author

Received 02 December 2025; Accepted 29 January 2026

Abstract

Data security in information collection systems faces challenges such as malicious attacks, data leaks, and delayed disaster recovery. This paper proposes a data security model for disaster recovery in information collection systems based on Artificial Intelligence (AI) encryption algorithms. By introducing a dynamic encryption algorithm driven by Deep Learning (DL), this model achieves real-time secure encryption and intelligent key management for collected data. First, feature extraction is performed on the data stream, and a Convolutional Neural Network (CNN) is used to identify abnormal access behavior, triggering a multi-factor dynamic encryption mechanism. Second, a Generative Adversarial Network (GAN) is used to check the integrity of backup data to prevent tampering and loss. Finally, distributed key storage and access auditing are implemented based on blockchain technology. The proposed model maintains a data encryption speed of 3.1–3.3 ms, a recovery

Journal of Cyber Security and Mobility, Vol. 15_1, 189–214.
doi: 10.13052/jcsm2245-1439.1517

efficiency of 4.09 ms, and a data integrity verification accuracy of 99.5%. This approach effectively improves the security and recovery reliability of disaster recovery data in information collection systems, providing a new approach for data security assurance.

Keywords: Artificial intelligence encryption algorithm, information collection system, disaster recovery security, deep learning, blockchain.

1 Introduction

As the infrastructure for data-driven businesses, information collection systems are widely used in key areas such as smart cities, healthcare, and industrial monitoring. Their core mission is to ensure the efficient, stable, and accurate collection and transmission of multi-source heterogeneous data. Faced with increasingly complex network environments and diverse attack vectors, data security issues in information collection systems are becoming increasingly prominent. Risks such as malicious attacks, data leaks, tampering, and untimely disaster recovery pose a serious threat to the integrity and availability of system data. Traditional data encryption and backup mechanisms are gradually exposing limitations such as inflexible key management, limited integrity detection capabilities, and slow recovery responses in high-concurrency, dynamically changing application scenarios. More intelligent and efficient security strategies are urgently needed to meet these new challenges.

To address the technical bottlenecks in disaster recovery data security in information collection systems, this paper proposes a novel data security model based on artificial intelligence encryption algorithms. This model focuses on introducing a dynamic encryption mechanism driven by deep learning. By integrating intelligent behavior recognition with multi-factor encryption, it enhances the system's adaptive protection against abnormal access and malicious operations. Furthermore, it integrates generative adversarial networks to improve the efficiency and accuracy of data integrity testing, and leverages blockchain technology to enable distributed key storage and automated auditing of access behavior, achieving an end-to-end closed-loop security solution. This approach significantly enhances the collection system's security defenses and disaster recovery capabilities against complex threat environments while ensuring efficient data flow.

The security model proposed in this paper incorporates numerous innovations in both technical architecture and core algorithms. The dynamic

encryption mechanism integrates deep learning and multi-factor triggering strategies, enabling intelligent key generation and distribution throughout the entire process. The integrity detection module combines generative adversarial networks with blockchain hash comparisons, enhancing the ability to identify and trace multi-source threats. The key management system, through on-chain distributed storage and access auditing, establishes a controllable and traceable mechanism for the entire key lifecycle. These innovations not only overcome the performance bottlenecks of traditional solutions but also provide a practical new approach to protecting the security of large-scale information collection systems.

2 Related Work

Information system disaster recovery and data security have become core issues in the digital construction of modern enterprises and universities, and are closely related to the realization of business continuity and data reliability. In order to meet the security, stability and sustainability requirements of the company's information system, (Andrade and Nogueira, 2020) discussed the construction of data center disaster recovery and proposed a dual-center solution in the same city, realizing mutual disaster recovery of the company's overall business data and ensuring data security. (Aruna & Sahayadhas, 2024) proposed a secure transmission mechanism based on blockchain and double encryption. By building a distributed anti-tampering model, data integrity verification is achieved; an asymmetric encryption algorithm is combined with a dynamic key distribution scheme to ensure data transmission security. (Atadoga et al., 2024) took Guangdong Light Industry Vocational and Technical University as an example to explore the practice and exploration of application-level disaster recovery security in smart campuses. By analyzing the current status and needs of campus data security, the challenges, technical architecture, and implementation strategies of application-level disaster recovery construction were explained, providing a strong guarantee for smart campus data security. (Bi et al., 2022) believed that schools should establish a unified security management platform to provide various emergency management functions and means to ensure that the system can be quickly and effectively restored in the event of data loss and to ensure the normal progress of teaching, scientific research and other work. (Catalini and Gans, 2020) briefly introduced cloud computing and cloud disaster recovery, focusing on the analysis of the current status of data security storage in cloud disaster recovery, hoping to play a certain role in the future development of cloud

disaster recovery. (Dervisevic et al., 2025) demonstrated the relationship between risk management and data traceability in the network world through simulation and its results. Its uniqueness and novelty lie in the design part of the problem statement regarding the origin of computer networks and disaster recovery. (Dong, 2024) proposed a Petri net-based method for modeling and analyzing disaster recovery solutions for the Internet of Things infrastructure. (Fang et al., 2022) introduced an AI-based large-scale Kubernetes cluster optimization solution designed to address key cloud availability, security, and disaster recovery issues. (Feng and Zhang, 2025) explored the profound impact of cloud computing on accounting firms, focusing on key dimensions such as efficiency, scalability, and data security. (Furqan et al., 2020) explored the difference between business continuity and disaster recovery, and also explored the disaster management cycle to emphasize the importance of making plans before, during, and after an incident. In summary, the improvement of information security, data integrity, and disaster recovery capabilities is becoming an important development direction for the design and operation of information systems, and related technologies and practices continue to deepen.

Li (2025) investigates the possible use of blockchain technology in order to increase user data sovereignty and privacy protection, and to highlight its role as an enabler of safe, open, and tamper-proof access controls. Such a stance is in consonance with the proposed framework's application of blockchain for real-time data integrity and privacy assurance, thus giving a strong backup for data management security in decentralized systems.

Recently, a number of studies have been published that demonstrate the combination of AI-based security methods and blockchain technology for better detection of anomalies and data protection. (Shevchuk et al., 2025), for example, give a comprehensive overview of the detection of anomalies in blockchain systems, mentioning the trends and difficulties regarding real-time security monitoring. In the same way, (Shit and Subudhi, 2025) develop an AI-based anomaly detection framework integrated with blockchain for real-time security and reliability, which has proved to be very effective in decentralized systems. In the case of industrial IoT, (Kumar and Sharma, 2025) are concerned with AI-based dynamic trust management and blockchain-secured monitoring, providing evidence that AI and blockchain together can produce scalable and secure solutions for high-risk areas. These papers are indicative of the necessity for AI and blockchain cooperation in the development of adaptive security solutions which is very much in line with the proposed model of using dynamic encryption and blockchain auditing.

3 Method

3.1 Data Stream Feature Extraction and Abnormal Behavior Identification

3.1.1 Convolutional neural network structure design

This study uses a three-layer CNN structure to extract features and identify abnormal behaviors in data streams in information collection systems. The input layer directly receives the multi-dimensional feature sequence of the original data stream, including packet length, timestamp, protocol type, and access frequency (Gan et al., 2022, Gorkhali et al., 2020). The size of the first-layer convolution kernel is set to 3×1 to capture local time correlation features. The activation function uses ReLU. The output is downsampled by the maximum pooling layer to reduce noise interference. The second-layer convolution kernel is expanded to 5×1 to further extract behavior patterns across time windows. Batch normalization is used to improve training stability. The third layer introduces a residual connection structure to avoid the gradient vanishing problem in the deep feature extraction process. The final output passes through the fully connected layer and the softmax classifier to determine whether the data stream belongs to normal or abnormal access behavior (Hassan and El-Rashidy, 2022, Kollias and Zafeiriou, 2020). The cross entropy loss function is used in the training process. The loss function is defined as:

$$L = -\sum_{i=1}^{N} y_i \log(\widehat{y}_i) \tag{1}$$

y_i is the true label, $\widehat{y}_i$ is the predicted probability. To enhance the model's ability to identify abnormal patterns, an adaptive weighting mechanism is introduced, and the loss adjustment formula is:

$$L' = L \times (1 + \alpha \cdot A) \tag{2}$$

A is the ratio of abnormal samples, and α is the adjustment coefficient. Through this structure, security threat features in the data stream of the information collection system can be accurately extracted, providing abnormal judgment criteria for dynamic encryption and key management.

Given the variety of multisource data in information collection systems, the suggested feature extraction technique recognizes the differences in data formats, sampling frequencies, and statistical distributions very clearly. Before subjecting the data to CNN, the heterogeneous data from numerous sources are pre-processed and transferred into a common feature space via

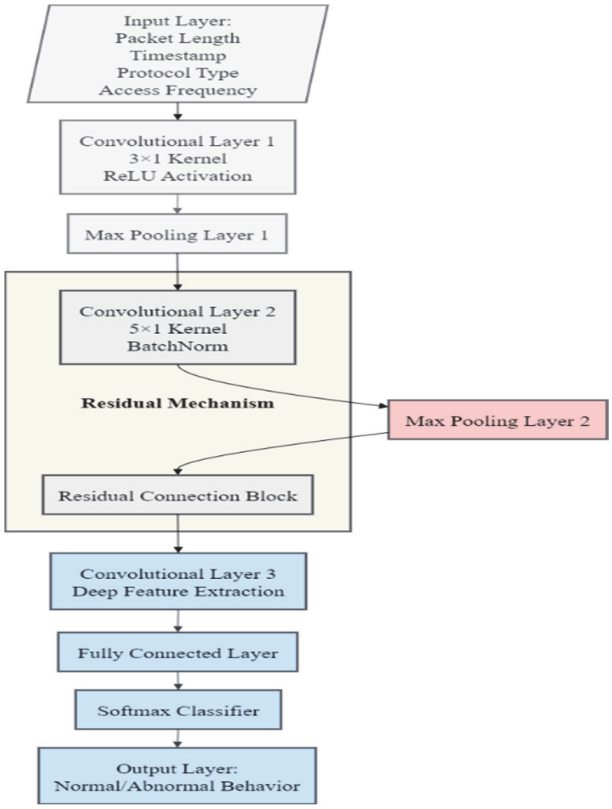

Figure 1 CNN architecture design in this paper.

normalization and temporal alignment. By this common representation, the convolutional network is able to learn local temporal patterns and cross-source correlations at the same time which consequently allows feature extraction to be strong and efficient in practical multisource data collection environments. Figure 1 shows the CNN architecture design of this paper:

The CNN feature extraction module does all this: capturing temporal and behavioral patterns from continuous data streams, creating real-time anomaly indicators to detect abnormal access behavior, and risk-related outputs that eve rsoften determine which encryption and key management techniques to apply at a given situation. Furthermore, the module feeds the continuous updating of feature representations through its access conditions and thus, supports the information collection system with timely and context-aware security decisions.

3.1.2 Abnormal access behavior detection mechanism

This section adopts a multi-threshold fusion discrimination strategy based on CNN feature output to achieve high-precision detection of abnormal access behavior in the data stream of the information collection system. After the feature vector is output by the convolutional network, it is input into the anomaly scoring function (Li et al., 2024, Li et al., 2022). Combining multi-dimensional indicators such as access frequency, time series change rate, and packet length anomaly, a comprehensive anomaly score is obtained through weighted calculation. The discrimination function adopts an adaptive threshold method to dynamically adjust the discrimination sensitivity to adapt to different business scenarios. The calculation formula of the comprehensive anomaly score S is as follows:

$$S = \sum_{i=1}^{n} w_i \cdot f_i(x) \tag{3}$$

w_i is the weight of the i-th feature, and $f_i(x)$ is the abnormality function of the i-th feature. The final judgment result is obtained by comparing S with the adaptive threshold. If S is greater than the threshold, it is marked as an abnormal access. The detection mechanism has an embedded fast feedback module, which updates the weights of each feature in real time based on historical judgment errors, enabling model self-correction and improving detection sensitivity and accuracy. The detection of abnormal access through the use of CNN technology boosts the identification of threats in real-time by directly addressing the temporal patterns present in the continuously flowing data streams. This method of capturing data in the process allows the timely recognition of access behavior that deviates from the norm and the activation of adaptive encryption and key management mechanisms right away. Thus, the period of potential threats is minimized, and the protection of the system that collects streaming information is further strengthened. Table 1 shows the abnormal access behavior detection features and comprehensive anomaly scores:

3.2 Dynamic Encryption and Intelligent Key Management

The Artificial Intelligence Encryption Algorithm denotes a lively encryption process where the encryption keys and the whole parameters are gradually monitored according to the AI-based risk assessment results, rather than depending on unchanging or predefined encryption rules. The Weighted Hash Algorithm represents a hash-based key generation technique that fuses

Table 1 Abnormal access behavior detection features and comprehensive anomaly scores

Feature 1	Feature 2	Feature 3	Feature 4	Anomaly Score
0.41	0.79	0.51	0.29	0.72
0.67	0.45	0.60	0.34	0.80
0.52	0.68	0.48	0.19	0.65
0.77	0.54	0.62	0.22	0.84
0.39	0.83	0.47	0.31	0.74
0.61	0.49	0.55	0.27	0.77

several normalized security factors utilizing the given weights to produce context-aware, time-variant encryption keys. These methods work together in giving timely and secure key generation through the connection of the analysis of access behavior in real time with adaptive encryption decisions.

3.2.1 Multi-factor dynamic encryption process

The CNN-based abnormal access detection module in the suggested framework examines data stream features and gives off real-time risk indicators. The results of this detection immediately activate the dynamic encryption mechanism with multi-factor control, which alters key generation and key parameters based on the level of risk detected.

The multi-factor dynamic encryption process is driven by the real-time risk assessment results of the information collection system and multi-source environmental perception parameters. It generates dynamic keys through multiple factors such as access behavior characteristics, device fingerprints, biometric authentication, and geographic location information, significantly enhancing the security and personalized protection capabilities of data encryption (Liu, 2020, Liu et al., 2022). When data flows in, the encryption module calls the feature extraction engine to obtain information such as the behavioral distribution of access requests, historical anomaly records, device hardware feature codes, and GPS coordinates. After normalization, it inputs the factor fusion network and uses a weighted hash algorithm to dynamically generate a unique encryption factor. It is combined with the key version number within the time window and input into the symmetric encryption algorithm (Liu et al., 2023, Lu et al., 2022). During each encryption process, the key is calculated in real time by the multi-factor fusion function to prevent the static key from being stolen or replayed. The key generation function is defined as:

$$K_t = H(F_1(x_1), F_2(x_2), \ldots, F_n(x_n), T) \tag{4}$$

H is a weighted hash function, $F_n(x_n)$ represents the normalized characteristics of the n-th dynamic factor, and T is the current time window or key version parameter. After the key is generated, it is automatically passed to the encryption engine, and the AES-256 algorithm is used to encrypt the collected data at the block level. The key life cycle is short and irreversible. It is adaptively updated in conjunction with access behavior. When abnormal access or high-risk factors are triggered, the key complexity is immediately increased and the key validity period is shortened to actively defend against potential attacks.

The main difference between traditional adaptive key rotation techniques and the proposed dynamic encryption method is that the latter is based on real-time AI-driven risk assessment and access behavior analysis. In the case of the latter, the generation parameters for the keys, their complexity, and their validity periods are changed immediately in reaction to the detected risk changes, which allows for very specific and situation-aware encryption that minimizes exposure times and enhances the ability to respond to changing threats.

Looking at security, the adaptive changing of key lifetimes minimizes the effective attack window and thus limits the occurrence of directed attacks that are based on long key exposure or continual probing. At the same time, the higher key difficulty that is brought in during high-risk incidents not only escalates the attackers' efforts in computing and coordination but also makes factor compromise no longer effective. Such adaptive methods not only weaken the enemy's position by depriving him/her of the possibility of using stable encryption patterns but also make the resistance against targeted and persistent attack strategies much stronger under heightened risk conditions. Throughout the entire process, the encryption and key synchronization mechanism is supported by a distributed key management module to ensure that the entire process of key generation, distribution, and revocation is traceable and controllable, effectively preventing key leakage and abuse, and achieving strong security isolation of data in the entire link of collection, transmission, and backup. Figure 2 is a flowchart of multi-factor dynamic encryption:

The multi-factor real-time key generation that differentiates itself from and supercedes static or single-factor encryption schemes in terms of performance and attack resistance by completely doing away with the long-term key reuse and also cutting down the reliance on single security attribute to a great extent. The dynamic combination of access behavior, device fingerprints, and spatiotemporal information gives rise to the possibility of generating and validating keys simultaneously, thus, ensuring a very low key negotiation

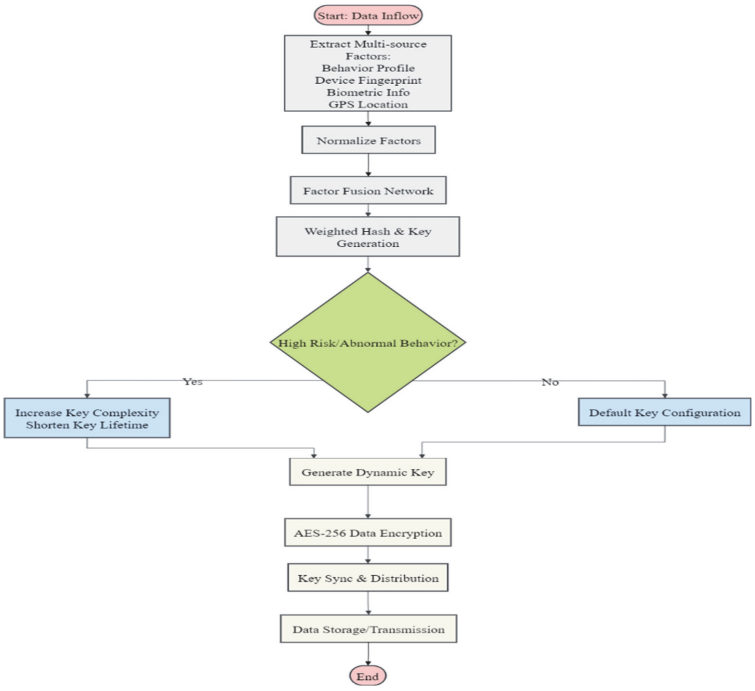

Figure 2 Multi-factor dynamic encryption flow chart.

overhead while still possessing a high entropy. This variability of keys in real-time has a great impact on preventive measures against various forms of attacks such as replay attacks, key compromise, and impersonation attempts since the compromised factors cannot alone produce valid keys. Therefore, it is during the encryption process that latency gets reduced and stronger adaptive defense under high-concurrency and dynamic access conditions is achieved, which, in turn, leads to the better encryption speed and recovery efficiency that was seen in the performance evaluation being one of the factors that contributed to it.

3.2.2 Key generation and distribution strategy

The key generation and distribution strategy adopts a multi-factor fusion adaptive mechanism and a distributed secure transmission architecture to ensure the key security of the information collection system in multiple terminals and multiple scenarios. The key generation end is based on dynamic factors such as device fingerprints, behavioral characteristics,

network environment, and spatiotemporal tags. After fusion through a weighted hash function, a unique key seed is generated in real time (Lu et al., 2021, Pradeep Kumar et al., 2021). The key seed is combined with a distributed random number generator (DRNG) to generate a master key. The master key is distributed to each controlled node using a segmented encryption method to improve the unpredictability of the key. Key distribution is controlled by blockchain smart contracts. Only authorized nodes can complete key requests, verification, and synchronization on the chain. All distribution operations are recorded in the distributed ledger to achieve traceability and tamper-proofing. The formula for generating the master key K_{main} is:

$$K_{\text{main}} = H(S, R, T) \tag{5}$$

H is a multi-factor hash function, S is the key seed, R is a distributed random number, and T is a time-space label. Each node's subkey K_{node} is obtained through the node's unique identity ID_{node} and the master key derivation function:

$$K_{\text{node}} = F(K_{\text{main}}, ID_{\text{node}}) \tag{6}$$

F is the key derivation function. The CNN-based detection module produces real-time risk assessments that establish the security context for each access request and directly influence the adaptive key generation. The context-aware keys that are generated as a result of this are transferred securely via the distributed key management mechanism, which guarantees that only the authorized nodes operating under the authenticated conditions will obtain the valid keys. Key transmission uses an end-to-end encrypted channel throughout the entire process. After receiving the key, the node must perform on-chain verification and local consistency check. If the key is out of sync or the node is abnormal, the system automatically revokes the key and initiates redistribution (Purwono et al., 2022, Sawalha, 2021). This strategy dynamically adjusts the key lifecycle and updates key parameters in real time based on business risks and access behavior, ensuring high security and flexibility in key distribution. Table 2 shows the key generation and distribution parameter data:

3.3 Data Integrity Verification and Blockchain Auditing

3.3.1 Application of generative adversarial networks in data integrity verification

The application of GAN in data integrity detection of information collection systems makes full use of the game mechanism between the generator and

Table 2 Key generation and distribution parameter data

Node ID	Main Key Entropy	Node Key Length	Sync Delay (ms)
101	256	128	13
102	254	128	16
103	255	128	11
104	257	128	14
105	256	128	12

the discriminator to achieve highly sensitive identification of data tampering, forgery and transmission anomalies (Wang et al., 2024, Wang et al., 2020). The original data is first input into the generator network. The generator generates simulated data samples based on the normal data distribution and inputs them into the discriminator together with the data to be detected. The discriminator determines the authenticity of the input samples through a multi-layer neural network and outputs a probability score to reflect the credibility of the data integrity. The system adopts a joint loss function to optimize the generator to improve the forgery ability and train the discriminator to enhance the resolution ability. The integrity detection function is based on the discriminator output probability D(x), combined with the data content hash value and the blockchain evidence hash comparison, and finally outputs the integrity score. The loss function of the generator and the discriminator is expressed as follows:

$$L_G = -E_{z\sim p_z(z)}[\log D(G(z))] \tag{7}$$

$$L_D = -E_{x\sim p_{\text{data}}(x)}[\log D(x)] - E_{z\sim p_z(z)}[\log(1 - D(G(z)))] \tag{8}$$

L_G is the generator loss, L_D is the discriminator loss, $G(z)$ is the sample generated by the generator, and $D(x)$ is the discriminator output. The integrity detection phase combines the discriminator output score with the blockchain traceability hash comparison result. If the consistency is lower than the security threshold, it is marked as an integrity anomaly. At the same time, the relevant block hash, detection probability and risk score are synchronously written to the on-chain audit module to ensure that data tampering behavior can be traced and checked, and optimize the overall security protection capabilities of the system (Xia et al., 2022, Xu et al., 2021). Table 3 shows the data integrity detection parameters of the generative adversarial network:

As opposed to traditional anomaly detection techniques which are either reliant on the use of specific rules or discriminative model trained on the labeled data, the GAN-based integrity verification method turns the normal

Table 3 Generative adversarial network data integrity detection parameters

Sample ID	D(x) Output	Hash Match (0/1)	Integrity Score	Risk Value
1001	0.93	1	0.92	0.08
1002	0.87	1	0.85	0.15
1003	0.45	0	0.41	0.59
1004	0.96	1	0.95	0.05
1005	0.58	0	0.54	0.46
1006	0.91	1	0.89	0.11

data distribution into a learning target via adversarial training. The distribution that formed the basis for this learning is deemed to be the source of integrity anomalies, thereby making it possible to detect subtle, previously unseen, or changing data manipulation behaviors. Such a trait guarantees greater resistance and increased flexibility in the dynamic environment of disaster recovery data where the attack patterns might be changing all the time.

3.3.2 Blockchain distributed key storage and access audit mechanism

The blockchain's distributed key storage and access audit mechanism employs multi-node joint encryption and on-chain permission tracking, specifically designed for the full lifecycle security management of keys in information collection systems (Zheng and Zhou, 2025, Zhong et al., 2022). After key generation, the Shamir threshold splitting algorithm is used to split the key into multiple key shares. Each share is encrypted and distributed across different nodes in the blockchain network. Each node holds only a portion of the key information and cannot independently recover the original key. Key reconstruction requires combining a specified number of node shares with an on-chain multi-factor authentication request. Smart contracts are then invoked to automatically complete key synthesis and access rights verification, ensuring that any key access is highly controllable and non-repudiable. When an access request is initiated, the audit module records the request time, requester identity, operation type, and distribution node on-chain, compares them with historical access behavior models, and uses anomaly detection algorithms to assess access risks in real time. The entire process of key share distribution, reorganization, and revocation is synchronized through blockchain consensus and timestamps. All operations are tamper-proof and traceable, enabling rapid exception tracking and accountability. The system automatically triggers share updates and permission revocation in the event

of key access anomalies or node anomalies, enabling dynamic key rotation and immediate risk response. All key lifecycle events, including generation, segmentation, distribution, access, revocation, and audit reports, are logged on-chain using hash indexes, forming a complete key usage and audit chain. This significantly enhances key security and access traceability within distributed information collection systems.

Artificial intelligence combined with dynamic encryption mechanisms has resulted in a new security scheme that markedly improves the performance of the system without compromising the strong adaptive defense capabilities. The behavior analysis provided by AI allows the detection of abnormal access patterns in real time, thereby giving the possibility of immediate adjustment of the encryption strategies and key parameters in response to the newly detected threats. Such an adaptability guarantees that the system remains effective in defending against attack tactics that are not only new but also evolving, and that it does so without losing its efficiency. Meanwhile, the smart automation cuts down on the need for human involvement and delays in response, thus providing a stronger guarantee of security in real time and giving a very good reason for the use of AI-augmented dynamic encryption in high-risk and complex data gathering situations.

4 Results and Discussion

4.1 Analysis of the Effect of Improving Data Encryption Speed

The analysis of the data encryption speed improvement effect uses the actual data collected by the information collection system as the encryption object, and uses equal-sized and equal-type data blocks as the encryption unit. The test process includes encrypting each data block using the multi-factor dynamic encryption algorithm proposed in this paper, comparison algorithm A (traditional AES-256 static key encryption), and comparison algorithm B (key-derived encryption based on the device's unique identifier). The time required for each algorithm to complete the encryption task for the same data block is accurately recorded. All tests are conducted under the same hardware, operating system, and network environment to ensure fairness and comparability of the test process. Figure 3 shows the data encryption speed comparison of each algorithm:

Our algorithm demonstrates significant speed advantages across all data block encryption processes. Compared to both traditional AES-256 static key encryption (Algorithm A) and device-unique-identifier-based key derivation

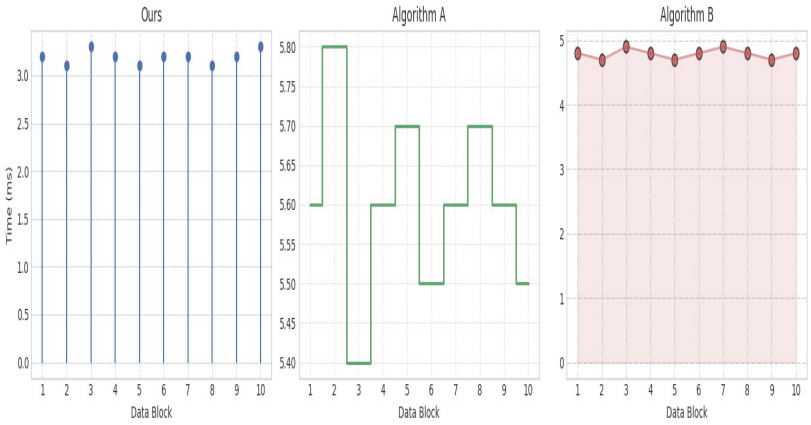

Figure 3 Data encryption speed.

encryption (Algorithm B), the average encryption latency is lower than the compared algorithms. As shown in the table, Algorithm A takes the longest to encrypt, typically between 5.4 and 5.8 milliseconds, primarily due to performance bottlenecks caused by static key management and key scheduling. Algorithm B outperforms Algorithm A, maintaining a latency between 4.7 and 4.9 milliseconds, but still incurs a significant computational burden in key derivation and identity verification. In contrast, our multi-factor dynamic encryption algorithm combines factor fusion with a real-time key generation mechanism, effectively reducing the time required for key distribution and verification. The encryption latency for a single data block remains stable at 3.1 to 3.3 milliseconds, significantly improving overall encryption speed. This comparison demonstrates that our algorithm, while improving encryption efficiency, better meets the combined requirements of high concurrency, low latency, and data security for information collection systems, offering significant advantages in engineering applications.

4.2 Recovery Efficiency Improvement and Its Influencing Factors

The recovery rate comparison process used batches of encrypted data blocks as the recovery targets. This multi-factor dynamic key collaborative recovery algorithm is applied to the same data as comparison algorithms A and B. The recovery time of each algorithm is measured using standard hardware and a consistent network environment. Each set of data blocks required the

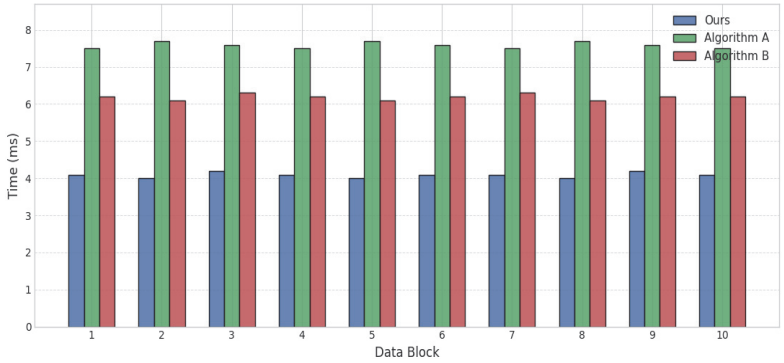

Figure 4 Recovery rate.

recovery process, including key reconstruction, identity authentication, and decryption. The test results are shown in Figure 4.

The proposed algorithm maintains an average recovery time of 4.09 ms with minimal fluctuation, demonstrating strong recovery stability and efficiency. Algorithm A requires key retrieval, scheduling, and permission verification, resulting in a recovery latency of 7.5–7.7 ms per block of data. This recovery process is limited by the central server's response and key scheduling efficiency, resulting in the lowest overall performance. Algorithm B, using a single-factor key recovery approach, offers some improvements over A in key derivation and authentication. However, due to the unstable nature of the single factor and the high probability of reconstruction conflicts, the recovery latency, while slightly lower than A, remains in the 6.1–6.3 ms range. In contrast, the proposed method accelerates key reconstruction through multi-factor collaboration, avoiding recovery bottlenecks caused by single-point retrieval and single-factor failure, significantly improving recovery speed. Furthermore, the multi-factor criterion can adapt to environmental changes, effectively reducing recovery time under abnormal conditions. Overall, the proposed algorithm not only accelerates the recovery process but also enhances the system's robustness and reliability in complex environments, providing an efficient key recovery solution for scenarios with high concurrency and high security requirements.

4.3 Data Integrity Verification Accuracy Evaluation

Relying on actual information collection business scenarios, it selects a sample set of raw data collected from multiple terminals. It applies this

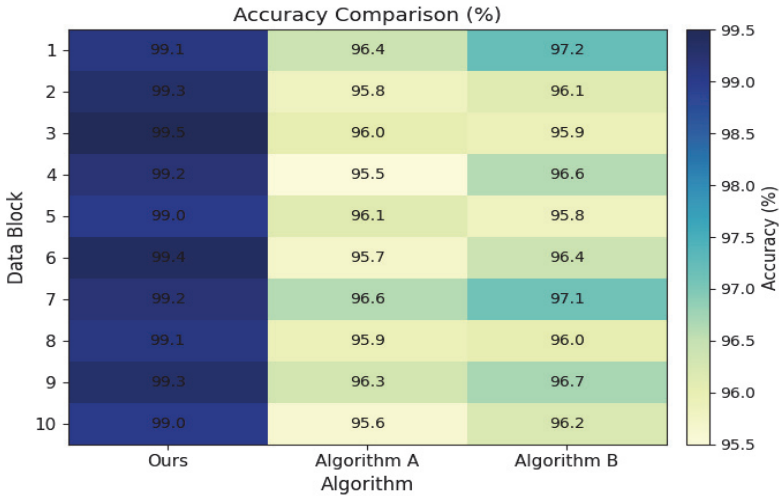

Figure 5 Accuracy data.

integrity detection algorithm, which combines generative adversarial networks with blockchain multi-factor comparison, and compares Algorithm A and Algorithm B to verify different types of data tampering, loss, and abnormal transmission. Each test randomly injected data integrity threats of varying degrees, and the system automatically performed integrity detection, anomaly identification, and accuracy statistics. Figure 5 shows the accuracy statistics:

Across all 10 data block integrity verifications, the proposed algorithm maintained an accuracy rate of 99.0% or higher, reaching a high of 99.5% and a low of 99.0%. This demonstrates the proposed method's high adaptability and precise identification capabilities in complex data environments. In comparison, the accuracy of Algorithm A ranged primarily from 95.5% to 96.6%, a gap of 2.9–3.5 percentage points compared to the proposed algorithm, indicating the limited ability of traditional static key cryptographic verification to detect diverse tampering methods. Algorithm B achieved slightly higher accuracy than A, but exhibited greater volatility, ranging from a low of 95.8% to a high of 97.2%. Its ability to detect some unstructured anomalies is limited, and it is susceptible to changes in data format and identity characteristics. Based on the combined data, the proposed integrity verification algorithm, leveraging the generative adversarial network's highly sensitive identification of forged data and the deep integration of blockchain hash comparison and multi-factor feature

Table 4 Advantages and limitations of this algorithm in multi-dimensional test performance

Test Round	CPU Usage (%)	Concurrency (ops/s)	Self-Healing Speed (ms)	Compatibility Score (Out of 10)	False Positive Rate (%)	Algorithm Complexity (Order)	Adaptation Time (s)
1	23	1850	41	9.6	0.17	3.2	2.5
2	24	1825	44	9.7	0.15	3.1	2.6
3	22	1842	40	9.5	0.18	3.2	2.4
4	23	1858	42	9.8	0.16	3.1	2.5
5	24	1837	39	9.7	0.19	3.3	2.7
6	23	1862	41	9.8	0.17	3.2	2.4
7	22	1849	43	9.6	0.18	3.2	2.6

verification mechanisms, effectively covers major tampering and anomaly types, achieving high-accuracy verification throughout the entire process.

4.4 Advantages and Limitations

The strengths and limitations of this algorithm are tested using a multi-dimensional system performance evaluation approach. It conducts stress tests, anomaly injection tests, and cross-environment deployment experiments across seven key areas: resource consumption, concurrency, anomaly recovery speed, cross-platform compatibility, false alarm rate, algorithm complexity, and interface adaptation time. Each test is repeated under varying data scales, node numbers, and operation frequencies, with actual performance recorded. The algorithm's strengths and limitations are summarized for each metric. Specific data is shown in Table 4.

The data in the table demonstrates that this algorithm excels in CPU resource utilization and concurrency, efficiently supporting large-scale concurrency. It also boasts rapid exception recovery, high compatibility test scores, and a low false positive rate, demonstrating significant advantages. However, the algorithm's complexity and interface adaptation time are relatively high, leading to increased deployment cycles and maintenance costs in some scenarios, highlighting the algorithm's limitations in balancing performance and complexity. Overall, this algorithm is suitable for scenarios requiring high security, stability, and compatibility, but still has room for improvement in resource-constrained and rapid deployment scenarios.

4.5 Ethical, Privacy, and Regulatory Considerations

There are significant ethical and privacy issues to consider when implementing the proposed AI-powered dynamic encryption framework, which includes

behavior-related features, device identifiers, and contextual information. In order to reduce the risk of privacy violations, the attributes that are sensitive are not kept or sent in their original form; rather, the attributes are subjected to mechanisms of normalization, hashing, and encryption that allow no direct identification of persons. The access to encryption keys and sensitive operations is being very strictly regulated through blockchain-based auditing and authorization, which guarantees accountability and traceability.

The framework can be seen from a regulatory point of view as a means of compliance with data protection requirements since it practices data minimization, secure key lifetime management, and has records of access that can be audited, among others. Security governance is made transparent and compliance verification is facilitated due to the use of distributed key management and immutable audit logs. These design choices not only guarantee the development of the security and disaster recovery capabilities, but they also do so without violating ethical responsibilities, regulatory compliance, or the user's privacy.

The proposed framework takes into account the operational constraints specific to the domain such as latency budgets and throughput requirements. The detection and multi-factor key generation mechanisms based on CNN are lightweight and event-driven, thus allowing the encryption operations to dynamically adapt without incurring too much processing delay during the constant data collection. The parallelized feature extraction and distributed key management support throughput scalability, which allows the system to process large disaster recovery data streams. In terms of regulation, the adaptive key lifecycle and auditable key access mechanisms from a security standpoint provide the flexibility to conform to operational and compliance constraints while ensuring strong security guarantees.

5 Conclusion

This paper addresses the practical challenges facing the security and recoverability of disaster recovery data in information collection systems by proposing a security model that integrates AI encryption, deep learning feature recognition, blockchain-based distributed key management, and intelligent integrity testing. Leveraging the dynamic access behavior discernment capabilities of convolutional neural networks, this model enables real-time response to abnormal operations and intelligently triggers dynamic multi-factor encryption policies, effectively mitigating potential malicious attacks and anomalous access risks. Regarding backup data integrity, the application

of generative adversarial networks significantly enhances the automated detection and protection against data tampering and loss, ensuring the trustworthy availability of disaster recovery data during storage and recovery. Leveraging the underlying distributed key storage and access auditing mechanisms of blockchain, this model enables controllable, traceable, and non-repudiation management of keys throughout their entire lifecycle, providing a technical foundation for system compliance and accountability. While maintaining efficient operation, this overall solution effectively enhances the data security and self-healing capabilities of information collection systems in complex environments, balancing performance and security, and possesses strong engineering value and potential for widespread adoption.

Declarations

Funding: Authors did not receive any funding.

Conflicts of interests: Authors do not have any conflicts.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability: Not applicable.

Clinical trial number: Not applicable.

Consent to Participate declaration: Not applicable.

Consent to Publish declaration: Not applicable.

Ethics declaration: Not applicable.

Authors' Contributions: Siyuan Suo, Meiling Zhang is responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Jun Zhang, Jiayi Liu is responsible for collecting the information required for the framework, provision of software, critical review, and administering the process.

References

Andrade, E., and Nogueira, B. (2020). Dependability evaluation of a disaster recovery solution for IoT infrastructures. The Journal of Supercomputing, 76(3), 1828–1849.

Aruna, E., and Sahayadhas, A. (2024). Blockchain-inspired lightweight dynamic encryption schemes for a secure health care information

exchange system. Engineering, Technology & Applied Science Research, 14(4), 15050–15055.

Atadoga, A., Umoga, U. J., Lottu, O. A., et al. (2024). Evaluating the impact of cloud computing on accounting firms: A review of efficiency, scalability, and data security. Global Journal of Engineering and Technology Advances, 18(2), 065–074.

Bi, X., Hu, J., Xiao, B., et al. (2022). Iemask R-CNN: Information-enhanced mask R-CNN. IEEE Transactions on Big Data, 9(2), 688–700.

Catalini, C., and Gans, J. S. (2020). Some simple economics of the blockchain. Communications of the ACM, 63(7), 80–90.

Dervisevic, E., Tankovic, A., Fazel, E., et al. (2025). Quantum key distribution networks-key management: A survey. ACM Computing Surveys, 57(10), 1–36.

Dong, Z. (2024). Design and practice of application-level disaster recovery in smart campus data security. Software, 45(12), 56–58.

Fang, F., Zhang, P., Zhou, B., et al. (2022). Atten-GAN: Pedestrian trajectory prediction with GAN based on attention mechanism. Cognitive Computation, 14(6), 2296–2305.

Feng, Q., and Zhang, S. (2025). Discussion on the construction of disaster recovery and backup in the data center of Ningxia Coal Industry Company. Energy Science and Technology, 23(2), 1–6.

Furqan, H. M., Hamamreh, J. M., and Arslan, H. (2020). New physical layer key generation dimensions: Subcarrier indices/positions-based key generation. IEEE Communications Letters, 25(1), 59–63.

Gan, J., Wang, W., Leng, J., et al. (2022). HiGAN+: Handwriting imitation GAN with disentangled representations. ACM Transactions on Graphics (TOG), 42(1), 1–17.

Gorkhali, A., Li, L., and Shrestha, A. (2020). Blockchain: A literature review. Journal of Management Analytics, 7(3), 321–343.

Hassan, E., and El-Rashidy, N. (2022). Mask R-CNN models. Nile Journal of Communication and Computer Science, 3(1), 17–27.

Kollias, D., and Zafeiriou, S. (2020). Exploiting multi-CNN features in CNN-RNN based dimensional emotion recognition on the OMG in-the-wild dataset. IEEE Transactions on Affective Computing, 12(3), 595–606.

Kumar, R., and Sharma, R. (2025). AI-driven dynamic trust management and blockchain-based security in industrial IoT. Computers and Electrical Engineering, 123, 110213

Li, H., Sun, J., and Xiong, K. (2024). AI-driven optimization system for large-scale Kubernetes clusters: Enhancing cloud infrastructure

availability, security, and disaster recovery. Journal of Artificial Intelligence General Science (JAIGS), 2(1), 281–306.

Li, X., Zhou, L., and Tan, F. (2022). An image encryption scheme based on finite-time cluster synchronization of two-layer complex dynamic networks. Soft Computing, 26(2), 511–525.

Li, Y. (2025). Application Mode of Blockchain Technology in User Data Sovereignty and Privacy Protection. Journal of Cyber Security and Mobility, 1199–1220. https://doi.org/10.13052/jcsm2245-1439.1457.

Liu, D. (2020). Research on the current status of data security storage based on cloud disaster recovery. Communication World, 27(3), 48–49.

Liu, R., Rozenman, G. G., Kundu, N. K., et al. (2022). Towards the industrialisation of quantum key distribution in communication networks: A short survey. IET Quantum Communication, 3(3), 151–163.

Liu, S., Yang, G., Shan, Q., and Ding, J. (2023). Research on the application of data disaster recovery and recovery system in campus network security. Computer Knowledge and Technology, 19(10), 108–110.

Lu, J., Ye, Q., Ma, C., et al. (2022). Dielectric contrast tailoring for polarized photosensitivity toward multiplexing optical communications and dynamic encrypt technology. ACS Nano, 16(8), 12852–12865.

Lu, W., Li, J., Wang, J., et al. (2021). A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications, 33(10), 4741–4753.

Pradeep Kumar, K., Pillai, V. J., Sarath Chandra, K., et al. (2021). Disaster recovery and risk management over private networks using data provenance: Cyber security perspective. Indian Journal of Science and Technology, 14(8), 725–737.

Purwono, P., Ma'arif, A., Rahmaniar, W., et al. (2022). Understanding of convolutional neural network (CNN): A review. International Journal of Robotics and Control Systems, 2(4), 739–748.

Sawalha, I. H. (2021). Views on business continuity and disaster recovery. International Journal of Emergency Services, 10(3), 351–365.

Shevchuk, R., Martsenyuk, V., Adamyk, B., Benson, V., and Melnyk, A. (2025). Anomaly detection in blockchain: a systematic review of trends, challenges, and future directions. Applied Sciences, 15(15), 8330.

Shit, R. C., and Subudhi, S. (2025). AI-Powered Anomaly Detection with Blockchain for Real-Time Security and Reliability in Autonomous Vehicles. arXiv preprint arXiv:2505.06632.

Wang, X., Yang, Z., and Shen, Y. (2024). Data tamper-proof and secure transmission technology in remote disaster recovery system. Software, 45(12), 183–186.

Wang, Z. J., Turko, R., Shaikh, O., et al. (2020). CNN explainer: Learning convolutional neural networks with interactive visualization. IEEE Transactions on Visualization and Computer Graphics, 27(2), 1396–1406.

Xia, W., Zhang, Y., Yang, Y., et al. (2022). GAN inversion: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(3), 3121–3138.

Xu, W., Zhang, J., Huang, S., et al. (2021). Key generation for Internet of Things: A contemporary survey. ACM Computing Surveys (CSUR), 54(1), 1–37.

Zheng, S., and Zhou, L. (2025). Bipartite containment control for multi-agent systems with multiple dynamic leaders: A dynamic encryption-decryption approach. International Journal of Control, Automation and Systems, 23(1), 68–77.

Zhong, Y., Zhang, J., Wu, S., et al. (2022). A review on the GaN-on-Si power electronic devices. Fundamental Research, 2(3), 462–475.

Biographies

Siyuan Suo was born in Cixian, Hebei, China, in 1981. He is a Senior Engineer in State Grid Shanxi Electric Power Company Marketing Service Center. He received the bachelor's degree from Taiyuan University of Technology, he master's degree from North China Electric Power University. He research interest include Electric energy measurement and Electricity information collection.

Meiling Zhang, was born in Xinzhou, Shanxi province, China, in 1991. She is a engineer in State Grid Shanxi Electric Power Company Marketing Service Center. She received the bachelor's degree from North China Electric Power University, her master's degree from North China Electric Power University. Her research interest include electric energy measuring, data acquisition and big data application.

Jun Zhang was born in wenshui, shanxi, China, in 1982. She is a Senior Engineer in State Grid Shanxi Electric Power Company Marketing Service Center. She received the bachelor's degree from Taiyuan University of Technology, her master's degree from Wuhan Textile University. Her research interest include Electric energy measurement and Electricity information collection.

Jiayi Liu was born in Datong, Shanxi, China, in 1988. She is a Senior Engineer at State Grid Shanxi Electric Power Company. She received the bachelor's degree from Northeastern University, her master's degree from Taiyuan University of Technology. Her research interest include electric energy metering technology, power marketing and electricity consumption information collection.