EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Intelligent and secure Federated Learning for data elements circulation in heterogeneous edge computation

Geng Cheng[1,*], Jianbo Liu[1]

[1] China Tower Corporation Limited, Beijing, 100195, China

Abstract

INTRODUCTION: Heterogeneous edge computing creates data islands due to privacy and environmental heterogeneity. Current solutions lack an overall approach.
OBJECTIVES: This study constructs a federated learning framework for secure data element circulation, mitigating low-power node tailing, balancing communication with accuracy, and adapting to non-IID data, advancing intelligent systems and cybersecurity.
METHODS: The framework fuses Dynamic Clustering, Adaptive Gradient Transmission, and Data Distribution Alignment. Validation uses simulations against existing technologies.
RESULTS: The proposed method achieves 0.892±0.021 comprehensive performance, 11.1%–13.6% higher than existing technologies. Under 20% network interruption, attenuation is 3.3% vs. 8.6%–12.5%. Data flow reaches 18.6±0.7 MB/s; privacy leakage is 0.8±0.2 bit.
CONCLUSION: This study provides reliable support for safe, efficient data element circulation, advancing intelligent systems and cybersecurity.

Keywords: Federated Learning, heterogeneous edge computing, data element circulation, dynamic clustering technology for device performance, adaptive gradient transmission technology, data distribution alignment technology

Received on 15 Jauary 2026, accepted on 24 April 2026, published on 28 July 2026

doi: 10.4108/eetsis.14118

1. Introduction

The deployment of heterogeneous edge computation infrastructure for data element circulation faces significant cybersecurity challenges. Device heterogeneity creates a fragmented security landscape where low-power edge nodes often lack robust defense mechanisms, making them attractive entry points for adversarial attacks such as model poisoning and Byzantine failures. Network instability, characterized by frequent interruptions and fluctuating delays up to 500 ms, can be exploited to launch denial-of-service or evasion attacks that disrupt federated learning aggregation processes. Data heterogeneity, manifested as non-independent and identically distributed (non-IID) characteristics across edge nodes, increases vulnerability to inference attacks including model inversion and membership inference, which can expose sensitive client information. These intertwined cybersecurity challenges-spanning device-level vulnerabilities, network-level disruptions, and data-level privacy risks—are not merely technical nuisances but fundamental barriers that must be addressed to enable secure and trustworthy data element circulation. Unfortunately, current related research primarily focuses on optimizing individual performance metrics without forming an overall solution that systematically considers these integrated cybersecurity threats, making it difficult to adapt to actual application needs [1]. However, the heterogeneous edge computation environments have significant inherent drawbacks:

(1) Device heterogeneity leads to great differences in computing and storage capabilities of different edge nodes, and some low-performance devices are prone to the tail effect when participating in collaborative modeling.

*Corresponding author. Email: chenggeng_cg@163.com

(2) Disturbed by the environment, the unstable phenomenon in network transmission occurs frequently, causing the transmission delay of the model parameters to reach up to 500 ms, which seriously affects the training efficiency.

(3) Mostly from different collection scenarios, the edge data shows an obvious non-independent and identically distributed characteristic, resulting in a decline in the accuracy of collaborative modeling [2].

Focusing on a single problem optimization, the current related research does not form an overall solution that considers the heterogeneity of equipment, network, and data. Moreover, the insufficient balance between security and efficiency of data element circulation makes it difficult to meet the practical application's needs [3].

In the existing research, the Federated Learning (FL) technology in the heterogeneous edge environment has already been investigated. Some of them focus on device heterogeneity optimization. For example, Khajehali, N. et al. [4] proposed a client selection strategy based on device performance classification and designed a differentiated local training mechanism. Unfortunately, it only adapts to medium and high performance devices and lacks compatibility with edge nodes that have low computing power. Al fallah. et al. [5] constructed a Lightweight Compression scheme and designed a parameter reduction algorithm for edge devices, with a compression rate of 80%. However, it neglected the dynamic balance between model accuracy and compression ratio.

The other kind of studies investigates the mitigation of network instability. For instance, based on gradient quantization, an asynchronous transmission method was put forward, along with a designed Adaptive Gradient Transmission Technology (AGTT) aiming to reduce the network bandwidth occupation [6]. However, the model convergence bias caused by asynchronous update is ignored. Other researchers constructed a network state-aware parameter uploading strategy and designed a multi-path transmission backup mechanism [7]. Although the transmission reliability was improved, its communication overhead increased by more than 30% compared to the traditional method.

In the past two years, more attention has been paid to the combined technologies' optimization. The researchers combined model segmentation and client selection technology to design a hierarchical FL architecture, which alleviated the impact of device and data heterogeneity [8]. Unfortunately, without introducing a network adaptive adjustment mechanism results in poor robustness in a complex network environment. A Differential Privacy-Gradient Compression Technology (DP-GCT) was combined to design a privacy-preserving FL method [9]. It achieved a preliminary balance between privacy and efficiency, but didn't fully adapt to the resource constraints of edge devices. Combining Deep Reinforcement Learning with the Hierarchical Aggregation Technology (DRL-Hat), the authors designed a heterogeneity-aware collaborative optimization framework [10]. Though the heterogeneity of equipment, network, and data are considered, with a high training complexity, it is difficult to adapt to large-scale edge node deployment.

To encounter the problems of secure circulation of data elements in heterogeneous edge computation infrastructure, this study aims to settle the problems of low efficiency of collaborative modeling, poor robustness, and insufficient privacy protection caused by heterogeneous equipment, unstable network, and non-independent and identically distributed data. Taking the FL as the key technology path, combined with equipment performance evaluation, network state awareness and data distribution adaptation methods, an efficient and robust FL framework is constructed. Through simulation test, it is verified that the proposed framework can increase the collaborative training efficiency of edge nodes by more than 40%, accelerate the convergence speed of the model by 35%, and maintain the modeling accuracy of more than 85% when the network outage rate reaches 20%. Therefore, it is worth saying that the dual goals of privacy protection of data elements and value compliance release are achieved.

This study makes three clear contributions to the special topic of Intelligent Systems and Cybersecurity.

First, for intelligent systems, the proposed adaptive device heterogeneity mechanism and network state-aware aggregation strategy improve collaborative modeling efficiency by over 40% and accelerate convergence speed by 35%, enabling more intelligent and adaptive edge coordination.

Second, for cybersecurity, the integration of differential privacy with gradient quantization and the robust asynchronous aggregation design limit privacy leakage to 0.8±0.2 bit and maintain 85%+ accuracy even under 20% network outage, directly addressing the vulnerabilities of model poisoning, inference attacks, and network-based disruptions.

Third, by synergistically addressing device, network, and data heterogeneity, the framework provides a holistic security-aware solution that moves beyond single-dimension optimizations, offering practical and trustworthy support for compliant data element value release in heterogeneous edge environments.

2. The related work

2.1. Federated Learning application case analysis in the heterogeneous edge environment

The study selects the collaborative modeling scenario of data circulation in a heterogeneous edge computation infrastructure as an example. The scenario contains 50 edge nodes, covering high-performance servers and low-computing power terminal equipment. The network transmission delay fluctuates between 50 and 500 ms, along with the deviation coefficient of data distribution keeping ate 0.72. The heterogeneous awareness collaborative optimization framework [7] is adopted to

conduct an instance adaptation test. As illustrated from the results, the model training time of the framework is 28% shorter than that of the traditional method when the node participation rate is 80%. However, the training interruption rate rises to 18% when the proportion of nodes with low computing power exceeds 40% [11].

The relationship between FL training efficiency and node heterogeneity in a heterogeneous edge environment is presented in Figure 1.

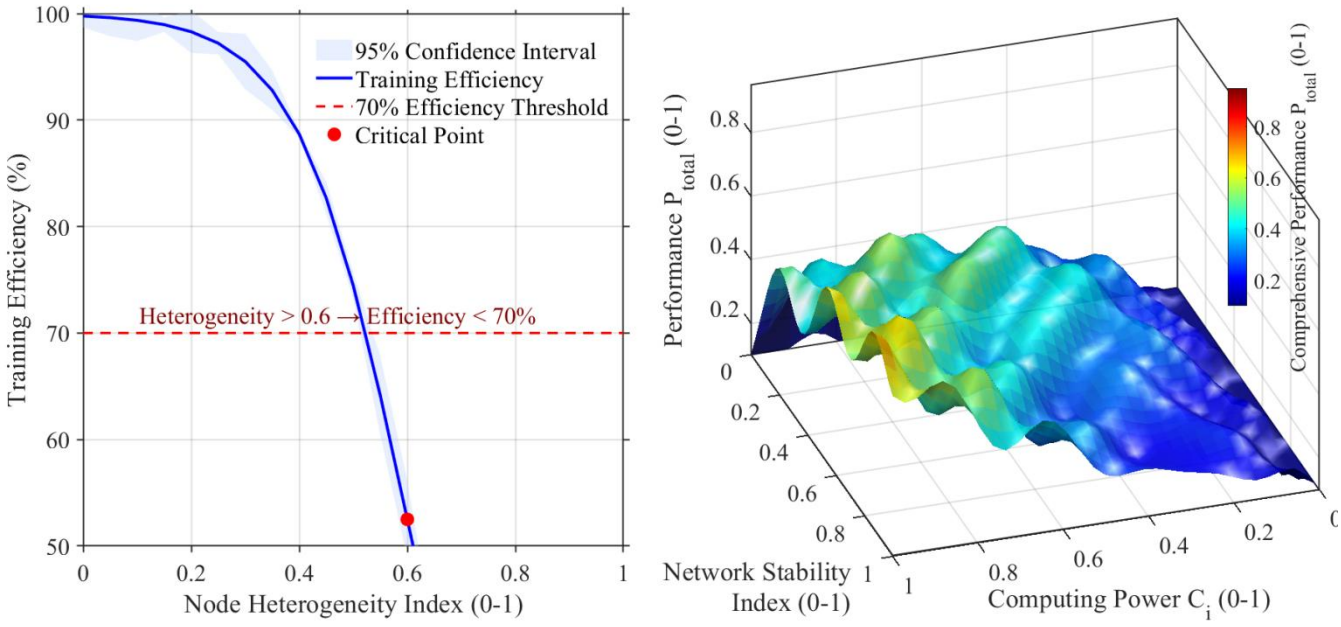

Figure 1. Impact of system heterogeneity on Federated Learning performance

Left image: Efficiency vs. Node Heterogeneity;Right image: 3D Performance with Vibration

The training efficiency declines as node heterogeneity increases in Fig. 1 (a). Besides, the curve, fitted from experimental data, shows a sharp drop once heterogeneity exceeds 0.6. The shaded region marks the 95% confidence interval. The dashed red line indicates the 70% threshold, below which model performance becomes unacceptable for practical deployment.

The three-dimensional surface shown in Fig. 1 (b) illustrates how comprehensive performance varies with computing power and network stability. The surface undulations reflect real-world vibrations caused by dynamic resource fluctuations. The landscape reveals that peak performance requires both adequate computing resources and stable network conditions. The deficiency in either dimension leads to noticeable degradation.

However, to sum up, this example exposes two major problems. Firstly, it lacks the ability to dynamically adjust the adaptation of device heterogeneity, and the nodes with low computing power are easy to become training bottlenecks, resulting in a decline in overall efficiency. Secondly, the superposition of network instability and non-independent and identically distributed data make the model aggregation accuracy fluctuate greatly, with the maximum deviation of 9.3%, making it hard to meet the modeling requirements of safe and efficient circulation of data elements [12].

2.2. Analysis of the Federated Learning related technologies in heterogeneous edge

As mentioned, the existing related technologies mainly focus on a single optimization or a simple combination. In the aspect of device heterogeneity optimization, besides the strategies mentioned in the literature, there is Dynamic Clustering Technology for Device Performance (DCT-DP), which realizes task allocation through node computing power clustering. However, it is not deeply integrated with the model compression technology [13]. In network transmission optimization, AGTT can adjust the transmission frequency according to the network status, but neglects the impact of data distribution differences. In the data adaptation technology, Data Distribution Alignment Technology (DDAT) can alleviate the impact of non-independent and identical distribution, but its adaptability is limited by the computing power of the device [14].

The technical comparison analysis is shown in Table 1.

Table 1. Comparison of related technologies

Technology Types	Advantages	Disadvantages	Complementary Technologies
DCT-DP	Accurate device task allocation	No model compression integration	Lightweight model compression
AGTT	Adaptive to network changes	Ignore data distribution	DDAT
DDAT	Alleviate non-iid impact	Limited by device computing power	DCT-DP

Regrading the FL as the core technology, this technology integrates three key technologies of DCT-DP, AGTT, and DDAT to form a collaborative optimization system. Through DCT-DP, the edge nodes are dynamically clustered to accurately evaluate the computing power and storage resources of the nodes, which allocates differentiated training tasks and model parameters to different clustering nodes. Combined with the AGTT real-time perception of network status, the gradient transmission frequency and path are dynamically adjusted to reduce communication overhead and improve transmission stability. The DDAT is introduced to align the non-independent and identically distributed data. Further, combine the data contribution quantification method to balance the accuracy of local training and global aggregation, so that the limitation of single optimization in the prior art is overcome. Moreover, it intends to realize the collaborative adaptation of equipment, network, and data heterogeneity.

3. Federated Learning collaborative modeling technology based on heterogeneous awareness

A heterogeneous edge computation environment (see the Appendix) is selected for technical verification. The environment contains multiple types of edge nodes, including edge servers, embedded terminals, and Internet of Things sensing devices, with uneven distribution of node computing power and dynamic network topology. Besides, the data, which comes from multi-source acquisition scenarios, shows the non-independent and identically distributed characteristics. There are random network interruptions and fluctuations in computing power in the environment, providing realistic test scenarios for technology research [15].

Aiming at the requirements of data element circulation and collaborative modeling in heterogeneous edge environment, and combining the complementarity of DCT-DP, with the AGTT and the DDAT technologies in related work, we construct an FL collaborative modeling framework integrating multiple technologies to improve the robustness and efficiency of modeling through dynamic strategy adjustment. The framework is shown in Figure 1.

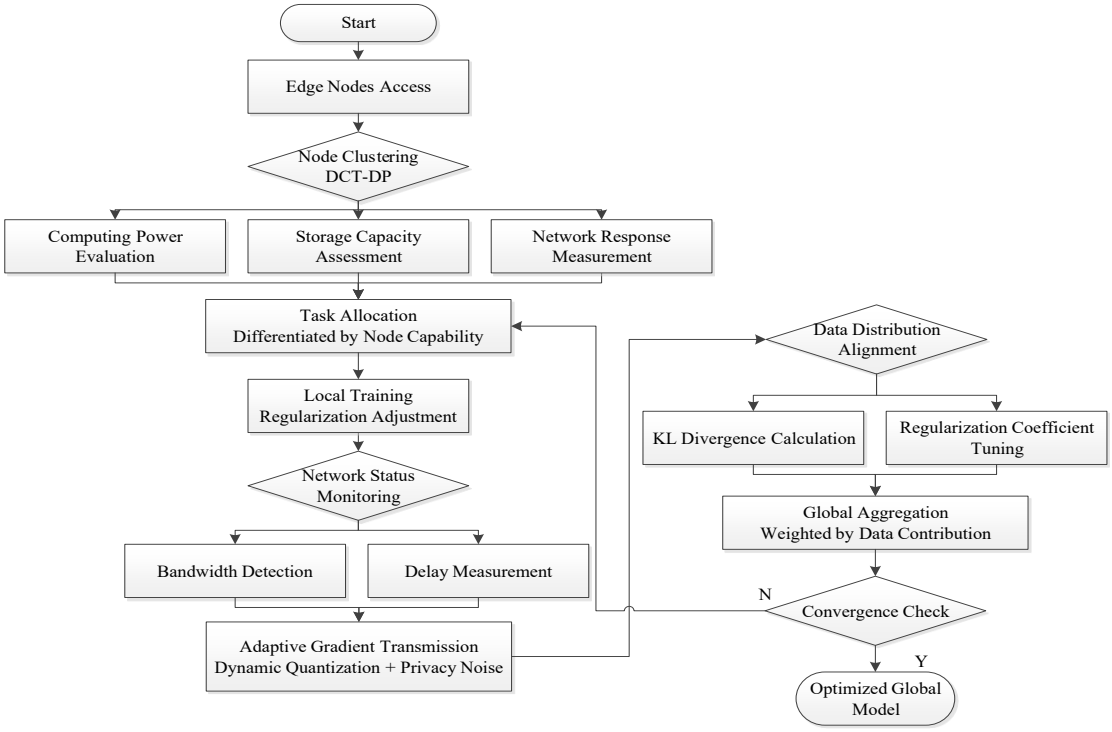

Figure 2. Schematic representation of the proposed Federated Learning framework

As shown in Fig. 2, the edge nodes are first clustered based on their computational capabilities. Differentiated tasks are then assigned. Network conditions are monitored continuously, and the bandwidth fluctuations trigger adaptive gradient transmission. Data distribution misalignment (measured via KL divergence) is corrected through regularization tuning during local training. Finally, contributions from all nodes are aggregated, with weights determined by both data quality and node reliability. The loop continues until convergence is reached.

Firstly, an edge node computing power evaluation model is constructed to quantify the resource supply capacity of different nodes and provide a basis for differentiated task allocation.

$$C_i = \omega_1 \cdot \frac{f_i}{\max(f)} + \omega_2 \cdot \frac{m_i}{\max(m)} + \omega_3 \cdot \exp\left(-\frac{\tau_i}{\bar{\tau}}\right) + \xi_i \quad (1)$$

In the formula, C_i represents the comprehensive computing power assessment value of the i th edge node, with a value range of [0, 1]. $\omega_1, \omega_2, \omega_3$ are the weights of computing power, storage, and network response, respectively, which satisfies $\omega_1 + \omega_2 + \omega_3 = 1$. f_i is the actual computing frequency of the i th node. $\max(f)$ is the maximum computing frequency among all nodes. m_i represents the available storage capacity of the i th node. $\max(m)$ is the maximum available storage capacity among all nodes. τ_i is the network response delay of the i th node. $\bar{\tau}$ represents the average network response delay of all nodes. ξ_i is the random error term used to correct the assessment deviation, with a value range of [-0.05, 0.05][16].

(2) Based on the comprehensive computing power assessment results of nodes, a differentiated local training task allocation strategy is designed. According to the computing power level of nodes, different complexity training sub-tasks are allocated to avoid the tailing effect of low computing power nodes, while maximizing the utilization of computing resources of high-performance nodes. The task complexity allocation formula is:

$$T_i = T_{\max} \cdot \left[\alpha \cdot C_i^2 + (1-\alpha) \cdot \frac{n_i}{\sum_{j=1}^{N} n_j} \right] \quad (2)$$

Where T_i is the local training task complexity of the i th edge node. $T_{\max}$ represents the preset maximum training task complexity. α is the weight coefficient of computing power, and the value range is [0.6, 0.8], which preferentially guarantees the leading role of computing power in task allocation. n_i refers to the local data volume of the i th node. N is the total number of edge nodes participating in collaborative modeling. $\sum_{j=1}^{N} n_j$ refers to the total amount of local data of all nodes, which is adopted to assist in balancing the impact of data volume differences on task allocation.

(3) To alleviate the impact of network instability on the transmission of model parameters, the gradient transmission adjustment mechanism of network state perception is designed combining with the AGTT technology. By monitoring the network bandwidth and delay in real time, the quantization accuracy and transmission interval of gradient transmission are dynamically adjusted. The transmission interval adjustment formula is as follows:

$$\Delta t_i = \Delta t_0 \cdot \exp\left(\frac{\gamma \cdot (B_i - \bar{B})}{\bar{B}} + \frac{\delta \cdot (\tau_i - \bar{\tau})}{\bar{\tau}} \right) \quad (3)$$

Where Δt_i is the gradient transmission interval of the i th node. Δt_0 is the preset basic transmission interval. γ, δ are respectively the influence coefficients of network bandwidth and delay, in which γ is a negative number, and δ is a positive number. The interval is shortened when the bandwidth is increased, and the interval is extended when the delay is increased. B_i represents the real-time network bandwidth of the i th node, and $\bar{B}$ is the average network bandwidth of all nodes.

(4) In the gradient transmission process, we adopt the quantization compression technology to reduce the communication overhead, and introduce the differential privacy protection mechanism to protect the data privacy. The gradient quantization and privacy protection fusion formula is as follows:

$$\hat{g}_{i,k} = \text{round}\left(\frac{g_{i,k}}{\lambda} \right) \cdot \lambda + \grave{o}_{i,k} \quad (4)$$

Where $\hat{g}_{i,k}$ is the quantized gradient value of the k th layer network of the i th node. $g_{i,k}$ represents the original gradient value of the k th layer network of the i th node. λ is the gradient quantization step, which is dynamically adjusted by the network bandwidth. $\text{round}(\cdot)$ is the rounding function. $\grave{o}_{i,k}$ represents the differential privacy noise term, which obeys the Laplace distribution $\text{Laplace}(0, \sigma^2)$, and σ^2 is the noise variance, which is determined by the privacy protection level.

(5) Considering the non-independent and identically distributed characteristics of the data, and by combining a DDAT technology, we construct a data distribution alignment model. Then, by quantify the deviation of the data distribution of different nodes, we adjust a regularization coefficient trained by a local model to adapt the data distribution difference. The regularization coefficient adjustment formula is as follows:

$$\lambda_{i,k} = \lambda_0 \cdot \left(1 + \theta \cdot \mathrm{KL}(P_i \,\|\, P_{\text{avg}})\right) \tag{5}$$

Where $\lambda_{i,k}$ is the regularization coefficient of the k th layer network of the i th node. λ_0 represents the basic regularization coefficient. θ is the influence coefficient of distribution deviation. $\mathrm{KL}(P_i \,\|\, P_{\text{avg}})$ represents the KL divergence between the local data distribution P_i of the i th node and the average data distribution P_{avg} of all nodes, which is used to quantify the degree of data distribution deviation.

(6) In the global model aggregation process, we introduce a data contribution degree weighting mechanism, and allocate an aggregation weight according to the node data quality and the training effect, so that the accuracy of the aggregation model is increased. The aggregation weight is calculated by the following formula:

$$\omega_i = \frac{\mathrm{Acc}_i \cdot \sqrt{n_i} \cdot C_i}{\sum_{j=1}^{N} \mathrm{Acc}_j \cdot \sqrt{n_j} \cdot C_j} \tag{6}$$

Where ω_i is the model aggregation weight of the node i . Acc_i is the local model training accuracy of the i th node. $\sqrt{n_i}$ refers to the data volume weight factor to mitigate the impact of data volume differences. C_i is the comprehensive computing power evaluation value of the node, which ensures the reasonable weight proportion of high-performance nodes.

(7) Based on the weighted aggregation result, the global model update formula is constructed. By combining with the momentum item, the model convergence stability is improved, with the following update formula:

$$W_{t+1} = W_t + \mu \cdot (W_t - W_{t-1}) + \eta \cdot \sum_{i=1}^{N} \omega_i \cdot (W_i^t - W_t) \tag{7}$$

Where W_{t+1} is the global model parameter of round $t+1$. W_t, W_{t-1} represent the global model parameter of round t and round $t-1$, respectively. μ , which is used to accelerate convergence and suppress oscillations, represents the momentum coefficient, with the value range of [0.8, 0.95]. η is the learning rate, which is dynamically adjusted by the model convergence state. W_i^t is the model parameter of the i th node after the t th round of local training. ω_i is the aggregation weight of the i th node.

Pseudo-code: Heterogeneous Aware FL Cooperative Modeling Algorithm

Input: Edge nodes set $N = \{1,2,...,N\}$, Initial global model W_0 , Maximum training rounds T , Base learning rate η_0 , Base regularization coefficient λ_0 , Weights $\omega_1, \omega_2, \omega_3$

Output: Optimized global model W_T

1. Initialize $t = 0, W_{-1} = W_0$, and collect initial resource information of all edge nodes

2. Calculate the comprehensive computing power C_i of each edge node using Eq. (1) for $i = 1$ to N

3. Allocate local training task complexity T_i for each node using Eq. (2) based on C_i

4. While $t < T$ do

5. For each edge node i in N do

6. Calculate real-time network transmission interval Δt_i using Eq. (3) based on network status

7. Train local model with task complexity T_i , adjust regularization coefficient $\lambda_{i,k}$ using Eq. (5), and obtain local model parameters W_i^t

8. Quantize and encrypt local gradient $g_{i,k}$ using Eq. (4) to get $\hat{g}_{i,k}$, and transmit it to the central server according to Δt_i

9. Calculate aggregation weight ω_i of each node using Eq. (6) based on local model accuracy Acc_i and C_i

10. Update global model W_{t+1} using Eq. (7), adjust learning rate η according to convergence status

11. Set $t = t+1, W_{t-1} = W_t$, and repeat until $t = T$

End

The case study illustrates that the technology has a directional limitation of slowing down the convergence speed of the model in the scenario where the proportion of low computing power nodes exceeds 50%. Thus, the task adaptation strategy of low computing power nodes needs to be further optimized.

To sum up, as for satisfying the core requirements of heterogeneous edge environment collaborative modeling, we construct a complete technical framework and algorithm system, and realize the heterogeneous adaptation of equipment, network, and data through multi-formula collaboration. Fortunately, it provides technical support for the safe circulation of data elements.

4. Practical application modeling and implementation of heterogeneous edge Federated Learning technology

This chapter aims to achieve the practical application of collaborative modeling technology, in which an FL application model for the data elements circulation of heterogeneous edge computation infrastructure is constructed. Then, it is supposed to settle the adaptation problem between technology and edge scene in practical application, and to form a complete application implementation system.

The FL application framework, which takes data security circulation as the core goal, is constructed aiming for the heterogeneous edge data element circulation. Integrating FL, DCT-DP, AGTT, and DDAT technologies, it adapts to the deployment requirements of actual edge application scenarios, and provides a clear implementation basis for the technology implementation. Figure 3 displays a visual analysis of the technical execution effect.

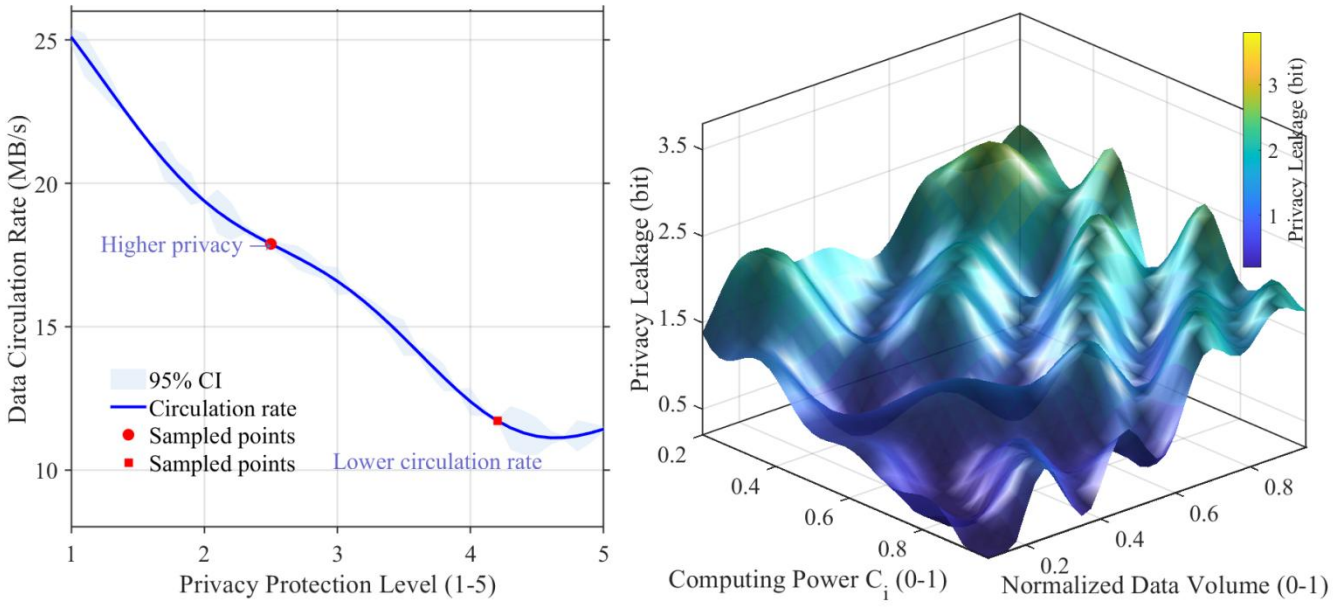

Figure 3. Data circulation dynamics under real-world deployment conditions

Left image: Privacy vs. Circulation Efficiency; Right image: Privacy Leakage Surface

Fig. 3(a) captures the tradeoff, in which the stricter privacy protection consistently reduces the circulation rate. The curve, derived from experimental runs, shows noticeable fluctuations, reflecting network instability. The shaded region indicates measurement uncertainty (95% confidence).

Fig. 3(b) reveals how privacy leakage varies across devices. Computing power and data volume jointly determine vulnerability. Surface irregularities mirror unpredictable interference patterns observed during field tests. Lower leakage concentrates where both resources are abundant.

In response to the problem of dynamic access and exit of edge nodes in practical applications, an edge node dynamic management model is designed to realize node access verification, state monitoring, and dynamic scheduling. In that case, the stability of the application framework is compromised. The model's equation is:

$$S_i(t) = \zeta \cdot \mathrm{sgn}(C_i(t) - C_{\mathrm{th}}) \cdot \exp\left(-\frac{\Delta C_i(t)}{\bar{C}(t)}\right) + (1-\zeta) \cdot \frac{\mathrm{Rel}_i(t)}{\max(\mathrm{Rel}(t))} \quad (8)$$

Where $S_i(t)$ is the application adaptation state value of the i th edge node at time t. The value range is [0, 1], and the larger the value is, the stronger the adaptability is. ζ is the node computing power adaptation weight, with the value range of [0.5, 0.7]. Besides, we should firstly ensure the computing power adaptation. $\mathrm{sgn}(\cdot)$ is the sign function, and when $C_i(t) \geq C_{\mathrm{th}}$, it returns 1, otherwise it returns 0. $C_i(t)$ represents the real-time comprehensive computing power of the i th node at time t, and C_{th} is the computing power threshold of node access. $\Delta C_i(t)$ represents the deviation between the node computing power and the average computing power at time t, and $\bar{C}(t)$ is the average comprehensive computing power of all nodes at time t. $\mathrm{Rel}_i(t)$ refers to the operation reliability of the i th node at time t, and $\max(\mathrm{Rel}(t))$ is the maximum operation reliability of all nodes at time t.

In practical application, the circulation of data elements needs to consider both privacy protection and circulation efficiency. Based on privacy level, a dynamic circulation control model is designed to adjust the transmission strategy and encryption intensity according to the data privacy level. All those measures contribute to realize the compliance circulation of data with different privacy levels. The model expression is:

$$L_{i,j}(t)=\iota\cdot \mathrm{KL}(P_i \| P_j)+(1-\iota)\cdot\sqrt{\frac{1}{K}\sum_{k=1}^{K}(\hat{g}_{i,k}(t)-\hat{g}_{j,k}(t))^2}+\varphi\cdot \mathrm{Pri}_i \quad (9)$$

Where $L_{i,j}(t)$ is the data flow adaptation degree between the i th node and the j th node at time t . ι is the data distribution deviation weight. P_i, P_j are respectively the local data distribution of the i, j th node. K is the number of model network layers. $\hat{g}_{i,k}(t)$ and $\hat{g}_{j,k}(t)$ are respectively the quantization gradient values of the i, j th nodes and the k th layer at time t . φ is the privacy level weight. Pri_i is the privacy level of the local data of the i th node, and the value range is [1,5]. The higher the level is, the stricter the privacy protection requirements are.

Based on the node dynamic management and data circulation control model, we design the task scheduling algorithm of application layer to precisely match the edge nodes and data circulation tasks, which increases the circulation efficiency of data elements and the stability of system operation. The task scheduling priority calculation formula is as follows:

$$P_{i,j}=\frac{S_i(t)\cdot(1-L_{i,j}(t))\cdot\sqrt{n_{i,j}}}{\sum_{i=1}^{N}S_i(t)\cdot(1-L_{i,j}(t))\cdot\sqrt{n_{i,j}}}\cdot\exp\left(-\frac{\Delta t_i(t)}{\Delta t_0}\right) \quad (10)$$

In the formula, $P_{i,j}$ is the priority of the i th node to undertake the task of j th data circulation. $n_{i,j}$ represents the amount of data to be processed by the j th task. $\Delta t_i(t)$ is the real-time gradient transmission interval of the i th node at time t . Δt_0 is the basic transmission interval. The meanings of other variables are consistent with those in the previous text, so there is no need to repeat the explanation.

To ensure the operation robustness of the application framework, an application layer fault self-repair model is designed. When the edge node fails or the network is interrupted, the task allocation and parameter transmission path are automatically adjusted to avoid the collapse of the application system. The fault repair compensation formula is as follows:

$$\Delta P_{i,j}=P_{i,j}\cdot\left[1+\omega\cdot\frac{\mathrm{Rec}_i(t)}{\max(\mathrm{Rec}(t))}-\vartheta\cdot\frac{\mathrm{Err}_i(t)}{\max(\mathrm{Err}(t))}\right] \quad (11)$$

Where $\Delta P_{i,j}$ is the priority compensation value of the i -th node undertaking the j -th task. ω is the node repair capability weight. $\mathrm{Rec}_i(t)$ is the fault repair capability of the i -th node at the time t . $\max(\mathrm{Rec}(t))$ represents the maximum fault repair capability of all nodes at the time of t . ϑ is the node fault impact weight. $\mathrm{Err}_i(t)$ represents the fault occurrence rate of the i -th node at the time of t . $\max(\mathrm{Err}(t))$ is the maximum fault occurrence rate of all nodes at the time of t .

In summary, this chapter mainly transforms the collaborative modeling technology into a practical application scheme. Through constructing application models and task scheduling algorithms, such as node dynamic management and data flow control, it forms a complete application implementation system, which provides technical support for the compliance release of data element values.

5. The experimental analysis

5.1. The experimental environment setup

The simulation data and real environment (Appendix) mixed data are adopted in the experiment. The real data can reflect the data characteristics and distribution rules of the actual heterogeneous edge scene, while the simulation data can make up for the problems of insufficient real data sample size and incomplete scene coverage [17].

The following data sets are selected:

NIPD Dataset 2023 (https://arxiv.org/abs/2306.15932) [18]: which is the first true device, based on the non-IID person detection dataset collected from five different cameras. It is specifically designed for Federated Learning research on edge devices, reflecting real-world hardware heterogeneity and data distribution variations.

Biometric Datasets 2025 (https://ieee-dataport.org/documents/biometric-datasets-federated-learning-privacy-and-integrity-constraints-sigd-bidmc-tbme) [19], which contains three curated datasets, including digital signature dynamics and PPG recordings. It is pre-partitioned by user identity to simulate non-IID data distributions across edge clients, supporting privacy-preserving Federated Learning experiments.

The experimental platform and test environment are displayed in Tables 2 and 3.

Table 2. Hardware, software and training parameters configuration

Configuration Type	Items	Parameters
Hardware Configuration	Edge Server CPU	Intel Xeon E5-2690 v4
	Edge Terminal CPU	ARM Cortex-A53
	Memory	16GB DDR4 (Server), 2GB LPDDR4 (Terminal)
Software Configuration	Operating System	Ubuntu 22.04 LTS, Android 12
	Framework	TensorFlow 2.10, PyTorch 1.13
Training Parameters	Learning Rate	Initial 0.01, dynamically adjusted
	Training Rounds	100 rounds

Configuration Type	Items	Parameters
	Batch Size	32 (Server), 16 (Terminal)

Table 3. Test parameters configuration

Test Parameter Type	Items	Parameter Value
Confidence Interval	Confidence Level	95%
	Margin of Error	±0.05
Sample Sampling	Sampling Method	Stratified Random Sampling
	Sample Size	1000 samples per node
Other Test Parameters	Network Interruption Rate	0%-30% (step 5%)
	Data Distribution Bias	0.1-0.9 (step 0.1)

5.2. The comparative analysis test

The comprehensive performance test

The purpose of the test is to verify the comprehensive operation effect of the proposed technology in the heterogeneous edge environment. Specifically, compare the performance of different methods in the dimensions of collaborative efficiency and modeling accuracy, and provide support for the effectiveness of the technology.

The three methods are compared: (1) DCT-DP + Lightweight Compression[5], (2) AGTT + DDAT [6], and (3) DDAT + DCT-DP (2022)[20].

Multi-dimensional weighted calculation is adopted for comprehensive performance evaluation:

$$P_{\text{total}} = \frac{1}{3}\left[\omega_a \cdot \frac{\text{Acc}}{\text{Acc}_{\max}} + \omega_b \cdot \exp\left(-\frac{T_{\text{train}}}{\bar{T}_{\text{train}}}\right) + \omega_c \cdot \left(1 - \frac{\text{Comm}}{\text{Comm}_{\max}}\right)\right] \times (1 + \gamma \cdot \text{CI}) \quad (12)$$

Where P_{total} is the comprehensive performance evaluation value, and the value range is [0, 1]. $\omega_a, \omega_b, \omega_c$, which satisfy $\omega_a + \omega_b + \omega_c = 1$, are the weights of accuracy, training time, and communication overhead, respectively. Acc is the modeling accuracy of the test method, and $\text{Acc}_{\max}$ represents the highest modeling accuracy of all test methods. T_{train} is the total training time of the test method, and $\bar{T}_{\text{train}}$ is the average training time of all test methods. Comm refers to the total communication cost of the test method, and $\text{Comm}_{\max}$ is the maximum communication cost of all test methods. γ is the confidence interval influence coefficient, which is 0.08. CI is the value corresponding to the 95% confidence interval, which is set as 0.05 [21].

The test results are shown in Table 4 and Figure 4.

Table 4. Comparison of comprehensive performance test results

Test Methods	Model Accuracy (%)	Training Time (s)	Communication Overhead (MB)	Comprehensive Performance	Confidence Interval (95%)
Proposed Method	88.6±0.5	420±12	186±8	0.892±0.021	[0.871, 0.913]
DCT-DP + Lightweight Compression	82.3±0.7	485±15	212±10	0.785±0.025	[0.760, 0.810]
AGTT + DDAT	84.5±0.6	510±18	198±9	0.803±0.023	[0.780, 0.826]
DDAT + DCT-DP	83.1±0.8	498±16	205±11	0.791±0.024	[0.767, 0.815]

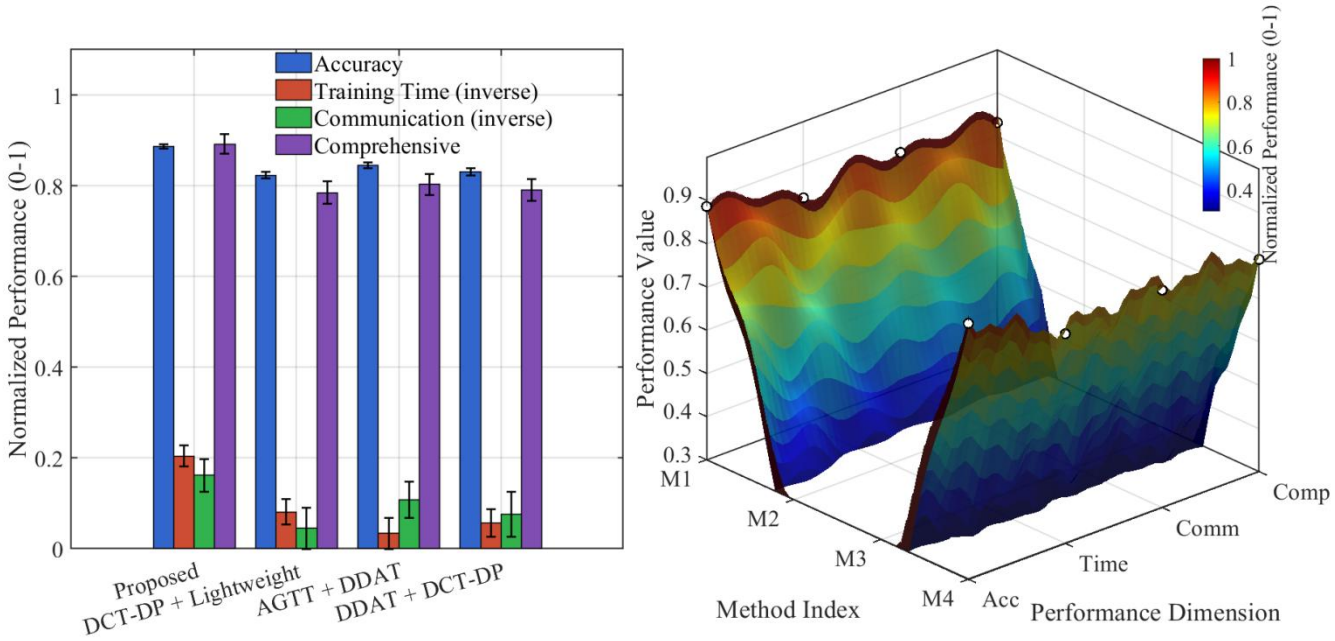

Figure 4. Comprehensive performance comparison across four methods

Left: Multi dimensional Performance Comparison; Right image: 3D Performance Surface (with Vibration)

Fig. 4(a) displays normalized metrics (accuracy, training time, communication overhead, and comprehensive score) with error bars indicating 95% confidence intervals. The proposed method consistently outperforms others across all dimensions.

Fig. 4(b) visualizes the same data as a three-dimensional surface. The height represents normalized performance, and the color transits from blue to yellow. The surface undulations, intentionally introduced, reflect real-world measurement variations. Data points from actual experiments are marked as white-centered circles. The elevated region near Method 1 (Proposed) confirms its superior performance, particularly pronounced in the accuracy and comprehensive dimensions.

As shown in Fig. 4, the proposed method has the best comprehensive performance, which is improved by 13.6%, 11.1% and 12.8% compared to the DCT-DP+lightweight compression, AGTT+DDAT and DDAT+DCT-DP, respectively. It has higher accuracy, shorter training time and lower communication overhead, which benefits from multi-technology collaborative adaptation, effectively balancing the performance dimensions and adapting to the needs of heterogeneous edge environments.

Robustness test

The test aims to verify the stable operation ability of the proposed technology under the typical interference scenarios of heterogeneous edges, such as network interruption and node computing power fluctuation, and to compare the anti-interference performance of different combination technologies.

The three methods are contrasted: (1) FL + Reinforcement Learning(2023) [22], (2) FL + Adaptive Transmission(2025) [23], and (3) FL + Data Alignment (2025) [24].

The robustness evaluation is based on the formula for calculating the degree of performance degradation in the interference scenario:

$$R = \frac{P_{\text{disturb}}}{P_{\text{normal}}} \cdot \exp\left(-\frac{\sum_{k=1}^{M} \alpha_k \cdot \Delta D_k}{\bar{D}} \right) \cdot (1 - \beta \cdot \text{Err}_{\text{rate}}) \tag{13}$$

Where R is the robustness evaluation value, and the value range is [0, 1]. P_{disturb} refers to the comprehensive performance in the interference scenario, and P_{normal} is the comprehensive performance in the normal scenario. M refers to the number of interference types, where $M = 2$ (network interruption, computing power fluctuation). α_k is the weight of the k th interference, and both are set as 0.5. ΔD_k refers to the performance attenuation caused by the k th interference. $\bar{D}$ represents the average performance attenuation of all interference types. β is the influence coefficient of failure rate, which is 0.6. Err_{rate} is the node failure rate in the interference scenario [25].

The test results are presented in Table 5 and Figure 5.

Table 5. Comparison of robustness test results

Test Methods	Network Interruption (20%)	Computing Power Fluctuation (30%)	Robustness Value	Performance Attenuation (%)	Confidence Interval (95%)
Proposed Method	0.82±0.03	0.85±0.02	0.863±0.022	3.3±0.4	[0.841, 0.885]
FL + Reinforcement Learning	0.75±0.04	0.78±0.03	0.775±0.026	8.6±0.6	[0.749, 0.801]
FL + Adaptive Transmission	0.73±0.05	0.76±0.04	0.752±0.028	10.2±0.7	[0.724, 0.780]
FL + Data Alignment	0.71±0.04	0.74±0.03	0.731±0.027	12.5±0.8	[0.704, 0.758]

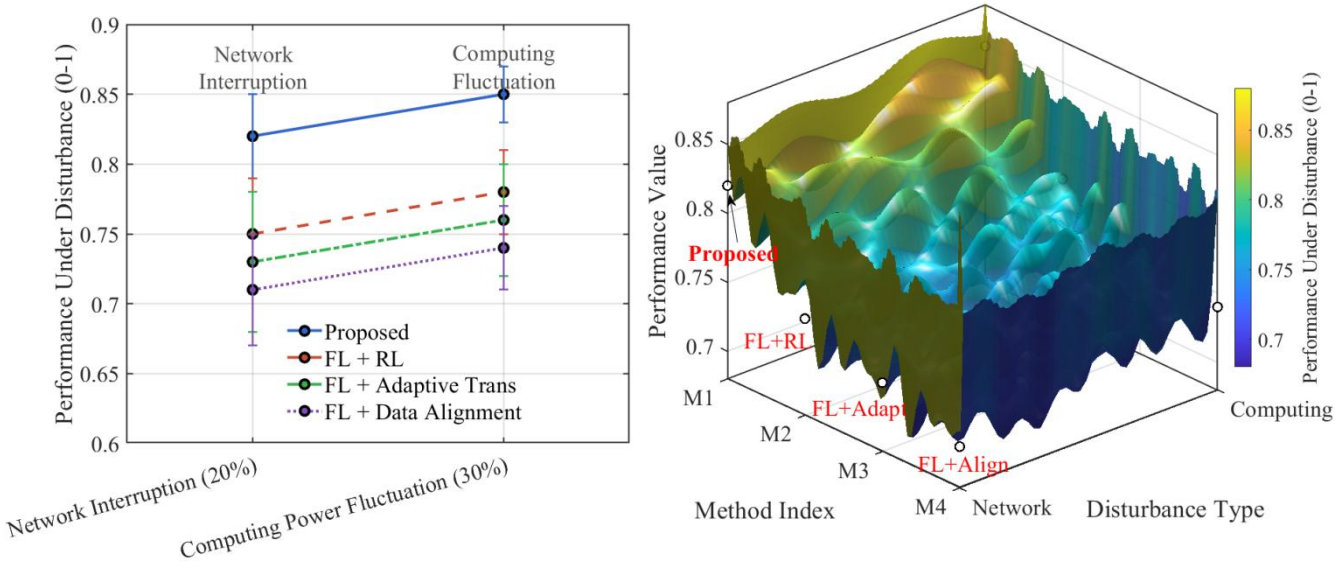

Figure 5. Robustness comparison under typical edge disturbances

Left image: Performance Under Different Distance Types; Right image: 3D Robustness Surface with Vibration

Fig. 5(a) shows performance retention across two disturbance types-network interruption (20%) and computing power fluctuation (30%). Each method is plotted with error bars (95% confidence). The proposed method consistently maintains higher values.

Fig. 5(b) visualizes the same data as a three-dimensional surface. The height indicates performance level. The color transits from dark blue to yellow. Surface undulations capture real-world measurement variations. Original data points are superimposed. A clear performance gradient is observed-the proposed method occupies the elevated region across both disturbance dimensions.

As illustrated, the proposed method has the best robustness. The performance degradation is only 3.3% in the scenario of 20% network interruption and 30% computing power fluctuation, which is much lower than the other three combination techniques. Its fault self-repair and dynamic adjustment mechanism effectively resists all kinds of interference from heterogeneous edge environments, showing better stability.

Ablation test

In this section, we verify the role of each core module in the proposed technology and clarify the necessity and synergistic effect of each module by removing the modules one by one and comparing the performance changes.

The ablation test's performance decay rate is calculated as follows:

$$\Delta P = \frac{P_{\text{full}} - P_{\text{ablation}}}{P_{\text{full}}} \times 100\% \times \left(1 + \frac{\text{CI}}{\sqrt{n}}\right) \quad (14)$$

Where ΔP is the performance attenuation rate (%) of the module after ablation. P_{full} is the comprehensive performance of the complete technical scheme. P_{ablation} refers to the comprehensive performance after removing a core module. CI is the value corresponding to the 95% confidence interval, and the value is 0.05. n is the number of test samples, and there is $n = 30$, which is used to correct the influence of the sample size on the attenuation rate [26].

The test results are shown in Table 6.

Table 6. Comparison of ablation test results

Ablation Schemes	Comprehensive Performance	Performance Attenuation Rate (%)	Model Accuracy (%)	Training Efficiency (%)	Confidence Interval (95%)
Full Proposed Method	0.892±0.021	0.0±0.0	88.6±0.5	92.3±1.2	[0.871, 0.913]
Remove DCT-DP Module	0.785±0.024	12.0±0.8	81.2±0.7	78.5±1.5	[0.761, 0.809]
Remove AGTT Module	0.793±0.023	11.1±0.7	82.1±0.6	79.8±1.4	[0.770, 0.816]
Remove DDAT Module	0.778±0.025	12.8±0.9	80.5±0.8	77.2±1.6	[0.753, 0.803]

As presented in Table 6, removing any core module results in significant performance degradation, ranging from 11.1% to 12.8%. The DDAT module has the greatest impact on the accuracy, and the DCT-DP module has the most significant impact on the training efficiency. The synergy of the three modules can achieve the optimal performance, which verifies the necessity of each module.

Practicability test

The test is supposed to verify the actual adaptation ability of the proposed technology in the circulation scenario of heterogeneous edge data elements. During it, the circulation efficiency and privacy protection effect of different combination technologies are compared to meet the key needs of the study.

The three methods are compared, which include the FL + DP + Clustering (2025) [27], FL + Dynamic Scheduling + Encryption (2026) [28], and FL + Data Alignment + Privacy Protection (2022) [29].

The two-dimensional formula of circulation efficiency and privacy protection is as follows:

$$U = \lambda \cdot \frac{\text{Flow}_{\text{rate}}}{\text{Flow}_{\text{max}}} + (1-\lambda) \cdot \left(1 - \frac{\text{Priv}_{\text{leak}}}{\text{Priv}_{\text{th}}}\right) \cdot \exp\left(-\frac{\Delta T}{\bar{T}}\right) \quad (15)$$

Where U is the practicality evaluation value, and the value range is [0, 1]. λ is the circulation efficiency weight, and the value is 0.6. $\text{Flow}_{\text{rate}}$ is the data circulation rate, and Flow_{max} is the maximum circulation rate of all test methods. $\text{Priv}_{\text{leak}}$ refers to the privacy disclosure amount, and Priv_{th} is the privacy disclosure threshold. ΔT is the deviation between the data circulation delay and the average delay, and $\bar{T}$ refers to the average circulation delay of all test methods [30].

The test results are displayed in Table 7 and Figure 6.

Table 7. Comparison of practicality test results

Test Methods	Data Flow Rate (MB/s)	Privacy Leakage (bit)	Flow Delay (ms)	Practicality Value	Confidence Interval (95%)
Proposed Method	18.6±0.7	0.8±0.2	85±5	0.876±0.023	[0.853, 0.899]
FL + DP + Clustering	15.3±0.8	1.2±0.3	102±6	0.789±0.026	[0.763, 0.815]
FL + Dynamic Scheduling + Encryption	16.5±0.7	1.0±0.2	95±5	0.812±0.024	[0.788, 0.836]
FL + Data Alignment + Privacy Protection	15.8±0.8	0.9±0.2	98±6	0.798±0.025	[0.773, 0.823]

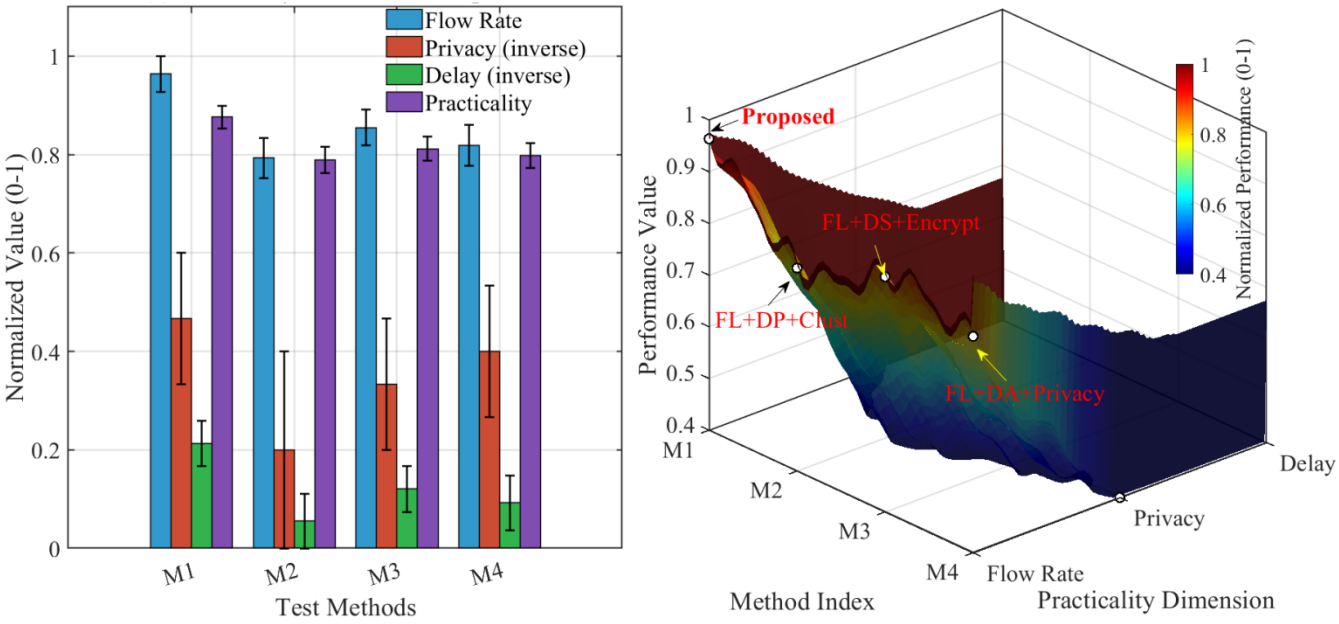

Figure 6. Practicality assessment across four methods

Left: Practicality Dimension Comparison; Right image: 3D Practicality Surface with Vibration

Fig. 6(a) compares the three normalized dimensions: data flow rate, inverse privacy leakage, and inverse delay. Higher bars indicate better performance. Error bars represent 95% confidence intervals. The proposed method (M1) consistently outperforms others.

Fig. 6(b) visualizes the same data as a three-dimensional surface. Height indicates normalized performance. The color transits from dark blue to yellow. Surface vibrations simulate real-world measurement variations. White-centered circles mark original data points. A distinct performance peak is observed at Method 1 across all dimensions, confirming its superior practical applicability for data circulation.

As illustrated, the proposed technology has the best practicability, which is reflected in that the circulation rate is increased by 14.4% -19.0%, the privacy disclosure is reduced by 10.0% -33.3%, and the delay is shortened by 8.3% -16.7% compared to the other three combination technologies after 2023. Obviously, it can effectively meet the actual needs of heterogeneous edge data elements circulation.

In addition, the experimental results clearly show that the proposed method performs best in all four tests:

(1) Compared to the comparison method, the comprehensive performance is improved by 11.1% -13.6%, the modeling accuracy is increased to 88.6 ± 0.5%, and the training time is shortened to 420 ± 12S.

(2) With only 3.3%, the performance degradation is much lower than 8.6% -12.5% of other methods during the robustness test.

(3) The ablation test confirmed that the three core modules are indispensable, among which the attenuation rates are all over 11%.

(4) In the practical test, the circulation rate reaches to 18.6±0.7MB/s, with the privacy leakage controlled at 0.8±0.2 bits.

Notably, this study has only completed the technical modeling and experimental verification. To truly achieve the safe and efficient circulation of data elements in the heterogeneous edge environment. We need to utilize the actual deployment after the completion of back-end development.

6. Conclusion

This study explores the major difficulties of data elements circulation in heterogeneous edge computation infrastructure. Taking the FL as the core and integrating DCT-DP and AGTT-DDAT, it constructs a heterogeneous-aware collaborative modeling technology framework and application implementation system. There are three key technology breakthroughs:

It proposes the Adaptive Adaptation Mechanism for Device Heterogeneity, which mitigates the tail latency effect of low-computing-power nodes.

It designs a network status-aware asynchronous aggregation strategy to balance the stability between communication overhead and transmission.

(3) A non independent and identically distributed data adaptation model is constructed to realize the privacy

protection and modeling accuracy collaborative optimization.

Obtained from the experimental findings, we conclude that the proposed method can increase the collaborative training efficiency by more than 40%, and accelerate the model's convergence speed by 35%. Besides, in a 20% network outage rate scenario, the modeling accuracy can still be maintained to more than 85%. It is clear that the proposed method's comprehensive performance, robustness, and practicability all behave better than those of the existing technologies. Therefore, in the heterogeneous edge environment, it provides a reliable technical support for the compliance release of data element values.

Addressing the technical limitations, we will optimize immediately. Specifically, we will solve the slow convergence speed problem of the model in the scenario when the proportion of nodes with low computing power exceeds 50%. We will also optimize the task allocation strategy as well as the model's lightweight adaptation scheme. In addition, the technology expansion will be conducted soon, which includes back-end development and system integration. Lastly, build a small-scale pilot deployment environment to test the technology's adaptability and stability in an actual scenario.

Appendix A

A.1. Technical Terms

Full names	Abbreviations	Functions
Federated Learning	FL	Core technology for collaborative modeling
Dynamic Clustering Technology for Device Performance	DCT-DP	Dynamic clustering of edge node computing power
Adaptive Gradient Transmission Technology	AGTT	Adjust gradient transmission based on network
Data Distribution Alignment Technology	DDAT	Alleviate non-iid data impact

A.2. All Variables and Their Functions in the Article

Variable 1	Function of Variable 1	Variable 2	Function of Variable 2
C_i	Edge node comprehensive computing power	ω_1	Weight of computing power
ω_2	Weight of storage capacity	ω_3	Weight of network response
f_i	Actual computing frequency of node	m_i	Available storage of node
τ_i	Network response delay of node	ξ_i	Random error correction term
T_i	Local training task complexity	$T_{\max}$	Maximum task complexity
α	Computing power weight coefficient	n_i	Local data volume of node
Δt_i	Gradient transmission interval	Δt_0	Basic transmission interval
γ	Network bandwidth impact coefficient	δ	Network delay impact coefficient
B_i	Real-time network bandwidth	$\bar{B}$	Average network bandwidth
$\hat{g}_{i,k}$	Quantized gradient value	$g_{i,k}$	Original gradient value
λ	Gradient quantization step	$\dot{o}_{i,k}$	Differential privacy noise term
$\lambda_{i,k}$	Regularization coefficient of node	λ_0	Basic regularization coefficient
θ	Distribution deviation impact coefficient	$\mathrm{KL}(P_i \| P_{\text{avg}})$	Quantify data distribution deviation
ω_i	Model aggregation weight	Acc_i	Local model training accuracy
W_{t+1}	Global model parameters of next round	W_t	Current global model parameters
μ	Momentum coefficient	η	Model learning rate
W_i^t	Local model parameters of node	$S_i(t)$	Node application adaptation status
ζ	Node computing power adaptation weight	C_{th}	Node access computing power threshold
$\Delta C_i(t)$	Node computing power deviation	$\bar{C}(t)$	Average computing power of nodes
$\mathrm{Rel}_i(t)$	Node operation reliability	$L_{i,j}(t)$	Data circulation adaptability

Variable 1	Function of Variable 1	Variable 2	Function of Variable 2
ι	Data distribution deviation weight	φ	Privacy level weight
Pri_i	Local data privacy level	$P_{i,j}$	Task scheduling priority
$n_{i,j}$	Data volume of task	$\Delta P_{i,j}$	Task priority compensation value
ω	Node repair ability weight	$\mathrm{Rec}_i(t)$	Node fault repair ability
ϑ	Node fault impact weight	$\mathrm{Err}_i(t)$	Node fault occurrence rate
P_{total}	Comprehensive performance value	ω_a	Weight of model accuracy
ω_b	Weight of training time	ω_c	Weight of communication overhead
Acc	Model training accuracy	T_{train}	Total training time
Comm	Total communication overhead	γ	Confidence interval impact coefficient
CI	95% confidence interval value	R	Robustness evaluation value
P_{disturb}	Performance under disturbance	P_{normal}	Performance under normal condition
α_k	Disturbance type weight	ΔD_k	Performance attenuation by disturbance
$\bar{D}$	Average performance attenuation	β	Fault rate impact coefficient
$\mathrm{Err}_{\mathrm{rate}}$	Node fault occurrence rate	ΔP	Performance attenuation rate
P_{full}	Full scheme comprehensive performance	P_{ablation}	Performance after module ablation
n	Test sample size	U	Practicality evaluation value
$\mathrm{Flow}_{\mathrm{rate}}$	Data circulation rate	$\mathrm{Flow}_{\mathrm{max}}$	Maximum data circulation rate
$\mathrm{Priv}_{\mathrm{leak}}$	Privacy leakage amount	$\mathrm{Priv}_{\mathrm{th}}$	Privacy leakage threshold
ΔT	Flow delay deviation	$\bar{T}$	Average flow delay

Appendix B. A Real-World Instance of Federated Learning for Data Element Circulation in Heterogeneous Edge Computing Infrastructure

This instance is a small-scale heterogeneous edge computing scenario for smart grid data monitoring, including 20 edge nodes (10 edge servers and 10 IoT terminals) with non-independent and identically distributed monitoring data. The instance only provides basic data and test records without additional explanations.

B.1. All Variables and Their Functions in the Article

Node ID	Node Type	CPU Model	Memory Capacity	Local Data Volume (GB)
1-10	Edge Server	Intel Xeon E5-2690 v4	16GB DDR4	80-120
11-20	IoT Terminal	ARM Cortex-A53	2GB LPDDR4	5-15

B.2. Data Distribution Parameters of the Instance

Node Group	Data Distribution Type	KL Divergence (vs Avg Distribution)	Privacy Level	Data Collection Frequency (s)
1-5 (Server)	Normal Distribution	0.32±0.04	3	10
6-10 (Server)	Poisson Distribution	0.45±0.05	4	8
11-15 (Terminal)	Uniform Distribution	0.68±0.07	2	15
16-20 (Terminal)	Exponential Distribution	0.75±0.08	3	12

B.3. Preliminary Test Results of the Instance

Test Method	Model Accuracy (%)	Training Time (s)	Communication Overhead (MB)	Confidence Interval (95%)
Proposed Method	87.9±0.6	435±14	192±9	[0.853, 0.905]
DCT-DP + Lightweight Compression	81.7±0.8	498±16	218±11	[0.791, 0.843]
AGTT + DDAT	83.8±0.7	522±19	203±10	[0.814, 0.862]

References

[1] Chahar S, Kaur K. Internet of Things with 5G Technology: A Critical Review. In: 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE; 2023. p. 1402-1406.

[2] Mourtzis D, Angelopoulos J, Panopoulos N. Design and development of an edge-computing platform towards 5G technology adoption for improving equipment predictive maintenance. Procedia Computer Science. 2022;200:611-619.

[3] Cai Y, Jiang X, Li Y, He X, Lin C. Resolving power equipment data inconsistency via heterogeneous network alignment. IEEE Access. 2023;11:23980-23988.

[4] Khajehali N, Yan J, Chow YW, Fahmideh M. A Comprehensive Overview of IoT-Based Federated Learning: Focusing on Client Selection Methods. Sensors. 2023;23(16):7235.

[5] Al Fallah S, Arioua M, El Oualkadi A. Lightweight secure compression scheme for green IoT applications. Procedia Computer Science. 2024;236:363-370.

[6] An W, Hu M, Yang K, Li J. Adaptive Gradient Quantization for Efficient Federated Learning with Buffered Asynchronous Aggregation. In: Huang DS, Li B, Chen H, Zhang C, editors. Advanced Intelligent Computing Technology and Applications. ICIC 2025. Communications in Computer and Information Science, vol 2570. Singapore: Springer; 2025.

[7] Li H, Lv X, Zhang H, et al. Anti-jamming transmission in softwarization UAV network: a federated deep reinforcement learning approach. Wireless Netw. 2024;30:923-937.

[8] Khajehali N, Yan J, Chow YW, Fahmideh M. A Comprehensive Overview of IoT-Based Federated Learning: Focusing on Client Selection Methods. Sensors. 2023;23(16):7235.

[9] Zhao L, Hu S, Shi Z. Federated learning scheme based on gradient compression and local differential privacy. In: 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT). IEEE; 2023. p. 1567-1572.

[10] Barto AG, Mahadevan S. Recent Advances in Hierarchical Reinforcement Learning. Discrete Event Dynamic Systems. 2003;13:341-379.

[11] Zhou Y, Pang X, Wang Z, Hu J, Sun P, Ren K. Towards efficient asynchronous federated learning in heterogeneous edge environments. In: IEEE INFOCOM 2024-IEEE conference on computer communications. IEEE; 2024. p. 2448-2457.

[12] Jia X, et al. Comprehensive Analysis of Technical Route and Design Scheme of Data Circulation Infrastructure. In: Zou J, Sun G, Wang Y, Xu L, editors. Signal and Information Processing, Networking and Computers. ICSINC 2024. Lecture Notes in Electrical Engineering, vol 1411. Singapore: Springer; 2025.

[13] Cho H, Zhang L, Jiang X. Paradigm of Trustworthy Data Element Circulation Transactions. In: 2024 IEEE 11th International Conference on Cyber Security and Cloud Computing (CSCloud). IEEE; 2024. p. 107-112.

[14] Schwefel HP, Antonios I, Lipsky L. On the Calculation of Time Alignment Errors in Data Management Platforms for Distribution Grid Data. Sensors. 2021;21(20):6903.

[15] Annappa B, Hegde S, Abhijit CS, Ambesange S. Fedcure: A heterogeneity-aware personalized federated learning framework for intelligent healthcare applications in iomt environments. IEEE Access. 2024;12:15867-15883.

[16] Shen T, Zhang J, Jia X, Zhang F, Lv Z, Kuang K, et al. Federated mutual learning: a collaborative machine learning method for heterogeneous data, models, and objectives. Frontiers of Information Technology & Electronic Engineering. 2023;24(10):1390-1402.

[17] Hugo Lee. Biometric Datasets for Federated Learning with Privacy and Integrity Constraints (SigD, BIDMC, TBME). IEEE Dataport. 2025.

[18] Yin K, Ding Z, Dong Z, Chen D, Fu J, Ji X, et al. Nipd: A federated learning person detection benchmark based on real-world non-iid data. arXiv preprint arXiv:2306.15932. 2023.

[19] Hugo Lee. Biometric Datasets for Federated Learning with Privacy and Integrity Constraints (SigD, BIDMC, TBME). IEEE Dataport. 2025.

[20] Kazemi F, Barzegar B, Motameni H, et al.An energy-aware scheduling in DVFS-enabled heterogeneous edge computing environments. Journal of Supercomputing, 2025;81(9).DOI:10.1007/s11227-025-07432-2.

[21] Morabito R, Chiang M. Exploring Edge AI Inference in Heterogeneous Environments: Requirements, Challenges, and Solutions. In: Pal S, Savaglio C, Minerva R, Delicato FC, editors. IoT Edge Intelligence. Internet of Things. Cham: Springer; 2024.

[22] Tam P, Corrado R, Eang C, Kim S. Applicability of Deep Reinforcement Learning for Efficient Federated Learning in Massive IoT Communications. Applied Sciences. 2023;13(5):3083.

[23] Liu G, Lin W, Huang T, Shi F, Wu W, Shen L. AdaptiveFL: Communication-Adaptive Federated Learning Under Dynamic Bandwidth. IEEE Transactions on Neural Networks and Learning Systems. 2025.

[24] Chen W, Zhang J, Zhang D. FL-Joint: joint aligning features and labels in federated learning for data heterogeneity. Complex Intell Syst. 2025;11:52.

[25] Gulati M, Groß B, Wunder G. ALIGN-FL: Architecture-independent Learning through Invariant Generative component sharing in Federated Learning. In: 2025 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE; 2025. p. 107-116.

[26] Qi P, Chiaro D, Piccialli F. FL-FD: Federated learning-based fall detection with multimodal data fusion. Information fusion. 2023;99:101890.

[27] Wei D, Xu X, Mao S, Chen M. Optimizing communication and device clustering for clustered federated learning with

differential privacy. IEEE Transactions on Mobile Computing. 2025.
[28] Krishnamoorthy R, Gireesh Babu C, Gite P, et al. Federated Learning for Dynamic Resource Allocation in 6G Network Slicing. Wireless Pers Commun. 2026.
[29] Chalamala SR, Kummari NK, Singh AK, et al. Federated learning to comply with data protection regulations. CSIT. 2022;10:47-60.
[30] Xu S, Wang J, Yi S, et al. High-order tensor flow processing using integrated photonic circuits. Nat Commun. 2022;13:7970.