Proceedings of the 5th International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/2026.31869

Selective Differential Privacy Federated Learning Framework Guided by Sensitivity

Xinge Huo

School of Artificial Intelligence and Computer Science, North China University of Technology, Beijing, China
2441973266@qq.com

Abstract. During the training process of federated learning, the model parameters uploaded by clients or the changes brought by parameter updates may still leak the relevant data information of the corresponding clients. As a mainstream privacy-preserving technology at present, differential privacy usually adopts the operation of adding uniform noise to parameters. However, this method often leads to slower convergence speed of model training, unsatisfactory model accuracy, and increased communication overhead. To address these issues, this paper proposes a dynamic differential privacy federated learning method guided by gradient-increment mask based on sensitivity. Starting from the update status of model parameters, this method establishes a sensitivity scoring mechanism according to the amplitude of parameter updates and gradient information, which is used to measure the degree of privacy leakage risk loss of different parameters in the training process. On this basis, a gradient-increment mask mechanism is added to selectively protect the parameters with high sensitivity scores, so as to reduce the unnecessary disturbance of noise distribution on low-sensitivity parameters. Furthermore, this paper constructs a dynamic privacy budget scheduling strategy guided by sensitivity, which flexibly adjusts the noise intensity according to the parameter sensitivity distribution to achieve refined and rational allocation as well as efficient utilization of privacy budget. Based on the analysis of these aspects, to resolve the contradiction between global noise injection and selective protection in differential privacy, this paper proposes a sensitivity-guided privacy budget allocation algorithm. This algorithm identifies the key parameters for model training through the mask consensus mechanism and allocates stronger noise protection to these parameters to meet the requirements of dynamic differential privacy. Theoretical analysis shows that this scheme satisfies the guarantee requirements of differential privacy. Experimental results indicate that under the same privacy budget, compared with the conventional full-coverage differential privacy federated learning, the model accuracy trained by this scheme in the same number of rounds is 4%-6% higher on several benchmark datasets, which basically proves the feasibility of the scheme proposed in this paper.

Keywords: Federated Learning, Differential Privacy, Mask, Sensitivity Scoring, Dynamic Privacy Budget

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.31869

1. Introduction

With the rapid evolution of artificial intelligence, people's demand for privacy protection is constantly increasing. Federated learning is a distributed machine learning method that does not require centralized training data. It realizes multi-party joint modeling through encrypted parameter exchange, ensuring that data can be trained without leaving local devices or affiliated institutions. Up to now, it has been applied in industries such as medical care, finance and the Internet of Things. By relying on the server to issue the initial model, and clients to upload parameters or parameter updates after completing local training, federated learning can reduce the risk of data leakage in this way. Conversely, recent studies have pointed out that the parameters uploaded by clients or their update content may still leak the individual data information of users, such as restoring the features of training samples through membership inference attacks or gradient inversion attacks, and even obtaining the private data information of clients locally. Therefore, how to further strengthen the intensity of privacy protection in federated learning is still a research topic of great concern.

Differential privacy is regarded as one of the potential and practical privacy-preserving technologies in the field of federated learning due to its rigorous mathematical privacy guarantee. At present, the common methods combining federated learning with differential privacy generally add uniform noise to all parameter updates uploaded by clients to protect personal data privacy. It should be noted that this kind of method of adding noise to parameters or their updates indiscriminately does not take into account the heterogeneous characteristics of model parameters in the machine learning process. Different parameters have obvious differences in the risk of privacy leakage and the impact on model performance. Adding noise of the same intensity to parameters during model training often leads to a significant decline in model accuracy, as well as a lot of unnecessary communication and computing overhead, thereby affecting the practical application effect of differential privacy in federated learning. In this regard, the practicality of such methods needs to be improved.

In recent years, many researchers have proposed many new mechanisms based on general differential privacy and adapted to different scenarios in the field of federated learning privacy protection, to meet the actual privacy protection needs of various scenarios. The theoretical basis of differential privacy can be traced back to the pioneering work of Dwork et al. [1], which provides a strict mathematical definition and provable privacy guarantee for data privacy protection. However, this kind of privacy guarantee causes excessive loss to the model. Traditional differential privacy mechanisms usually adopt a uniform noise injection strategy, that is, adding noise of the same intensity to the model parameters to be uploaded. This method aims at the security protection of privacy and does not conduct in-depth exploration on model accuracy and other aspects. Future research needs to make more trade-offs between accuracy and privacy [2,3]. After this research, several scholars including Liu proposed an adaptive differential privacy algorithm [4]. By analyzing gradient information to adjust the noise level at any time, this algorithm improves the model performance under the premise of using the same privacy budget. Although this approach has indeed found a way to balance privacy protection and model accuracy, the application scenarios of federated learning are becoming more and more extensive at present. Many users have higher requirements for the communication efficiency of the algorithm, hoping that the algorithm can save unnecessary overhead in the communication link. This adaptive differential privacy mechanism indeed has the problem of high computing and communication overhead [5], and the algorithm proposed by Liu still has much room for adjustment. To further improve the channel exchange efficiency, Zhang et al. constructed a sparse differential privacy algorithm [6]. By only injecting perturbations into some relatively important parameters, this algorithm significantly reduces the

DOI: 10.54254/2753-8818/2026.31869

computing and communication costs while ensuring privacy protection. Nevertheless, the sparse differential privacy algorithm lacks an effective parameter selection mechanism, making it difficult to accurately identify the most critical parameters for privacy protection. The inability to precisely select sensitive parameters leads to insufficient privacy protection intensity in some scenarios [7]. To solve this problem, several scholars including Wang proposed a gradient-guided differential privacy algorithm [8]. By analyzing the sensitivity information of model gradients, this algorithm can intelligently select the parameter subset that needs enhanced protection, and provide provable privacy guarantee while maintaining high model accuracy. However, it should be noted that the gradient-guided differential privacy algorithm is highly dependent on the accuracy of gradient information. If the gradient fluctuates sharply in a small range under the influence of noise that has been added, it will affect the calculation of the algorithm itself. Therefore, its application effect is relatively limited in some complex deep learning tasks [9].

To solve the intractable dilemma of difficult trade-off between privacy and accuracy, this paper focuses on the balance problem between privacy protection and model performance in federated learning, and constructs a sensitivity-guided selective differential privacy operation mode. This operation mode integrates parameter update amplitude and gradient information to build a parameter-level sensitivity quantification system, which accurately evaluates the privacy leakage risk of each dimension; designs a gradient-increment mask according to the quantification adaptability, which can implement selective noise injection on high-sensitivity parameters according to the mask, avoiding excessive disturbance to low-sensitivity parameters; further constructs a dynamic privacy budget scheduling strategy, which realizes the adaptive allocation of perturbation intensity according to the sensitivity distribution, improving the utilization efficiency of privacy budget. Theoretical analysis verifies that this mechanism satisfies the differential privacy guarantee. Experimental results show that under the same privacy budget, the model accuracy of the proposed method on multiple benchmark datasets is 4%-6% higher than that of the global noise addition strategy, and achieves considerable optimization in terms of communication overhead.

The main contributions of this paper can be summarized as follows:

1. It analyzes a sensitivity scoring mechanism based on parameter variation and gradient information, providing a new parameter screening mechanism for federated learning;
2. It designs a mask-driven selective differential privacy protection method, which applies perturbations to high-sensitivity parameters and adds a small amount of noise or no noise to low-sensitivity parameters. In this regard, a new way to balance accuracy and privacy is proposed;
3. Systematic experimental verification shows that the proposed method has good applicability in the non-independent and identically distributed federated learning scenarios, providing solid experimental basis for research in this field.

2. Introduction to related algorithms

The sensitivity quantification algorithm detects the parameters in each round of training process, and calculates a sensitivity quantification value through the relevant calculation methods of gradient and parameter variation. Its core idea is not to add noise of the same intensity to all model updates, but to allocate privacy budget differently according to the "sensitivity" of model parameters or data points, that is, their importance to the final model, so as to achieve a better trade-off between privacy and utility.

Traditional differential privacy federated learning mostly adds Gaussian noise of the same intensity to the model updates of each client in training. Although this framework is simple to use, it does not take into account the differences in the contribution of different parameters or different data

DOI: 10.54254/2753-8818/2026.31869

points to model training. Selective differential privacy designed based on sensitivity is built on the following key observations:

A. Sensitivity heterogeneity: In complex machine learning models, different parameters, that is, the proportion of the maximum parameter change caused by the modification of a single data point to the model weight, are different. Parameters that have a great impact on the final prediction results, that is, high-sensitivity parameters, need stronger privacy protection.

B. Optimized allocation of privacy budget: Allocate limited privacy budget to "where it is most needed": allocate more privacy budget to high-sensitivity parameters, that is, add less noise, and allocate less budget to low-sensitivity parameters, that is, add more noise. In this way, while maintaining the overall privacy protection level, the training utility of the model is maximized.

The selective differential privacy federated learning algorithm designed based on sensitivity represents the development trend of privacy-preserving machine learning towards more refined and intelligent direction.

3. Research methods in this paper

3.1. Federated learning

This study constructs a complete federated learning cycle integrating sensitivity adjustment mechanism. The process starts with task initiation, and clients first carry out initialization, including loading the model structure and setting the privacy level. Then the central server broadcasts the current global model to all participating clients. After receiving the model, the client performs local model training with the goal of optimizing the loss function. After completing the calculation, the process enters its core privacy protection link—noise injection with sensitivity adjustment, that is, applying differential perturbation according to the sensitivity of model updates to achieve differential privacy. After that, the client encrypts the processed parameters and transmits them back to the server through a secure channel. After the server collects all updates, it integrates these parameters by using a robust aggregation strategy to prevent possible malicious or low-quality updates, so as to obtain a new generation of global model. At this point, the system will check whether the performance of the new model meets the standard: if not, a new round of cycle will be started, and the new model will be issued from the server again; if it meets the standard, the whole federated learning task will be completed and ended.

DOI: 10.54254/2753-8818/2026.31869

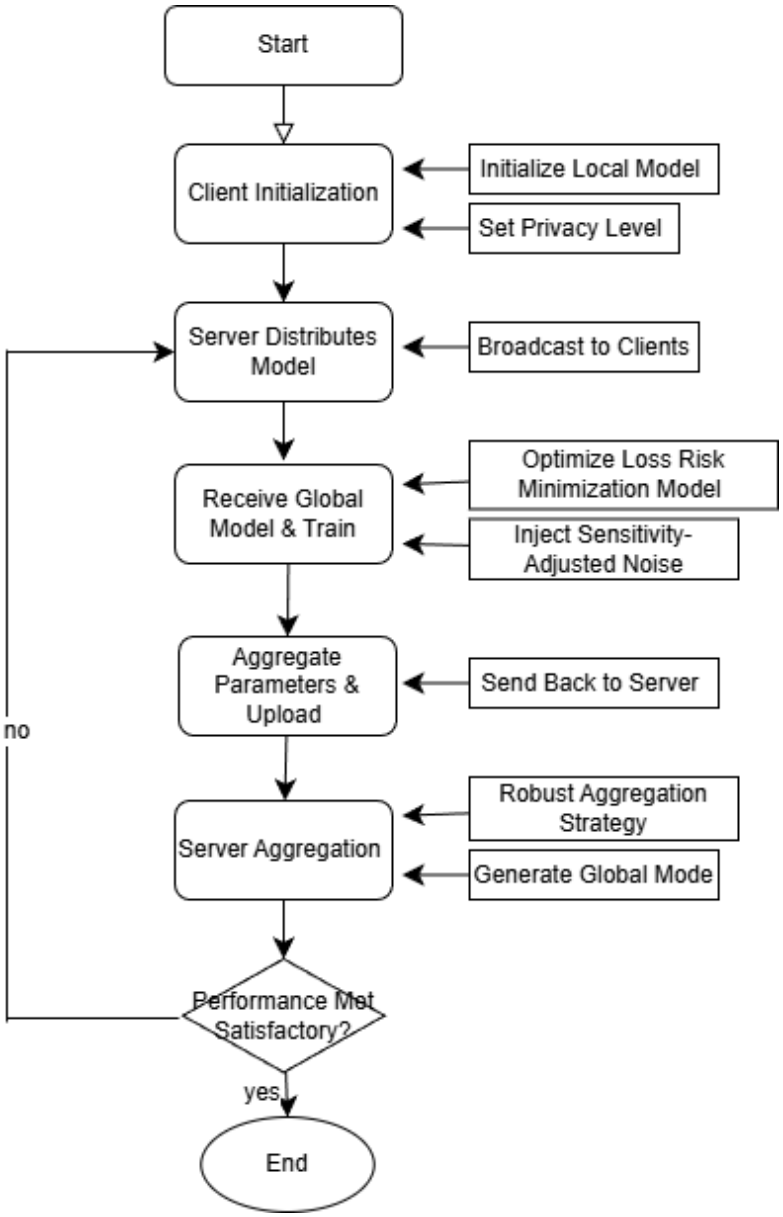

Figure 1. Flowchart of federated learning

There are N clients C training the model under the overall planning of the central server S. The client will use the local data sample D to train the global model parameter w issued by the central server. The core goal of federated learning is to collaboratively train a global model without sharing local data. Its loss risk minimization problem can be expressed as:

$$\min_{w} F(w) = \frac{1}{N} \sum_{i=1}^{N} F_i\left(w; \mathscr{D}\right) \tag{1}$$

Among them, F_i is the local risk related to model training in the i-th client, which can be specifically expressed as:

$$F_i(w) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathrm{L}\left(w; \rho_j, \varepsilon_j\right) \tag{2}$$

DOI: 10.54254/2753-8818/2026.31869

L is the loss function, whose significance is to calculate the average loss of all samples local to the client. Based on the analysis here, the loss function can preliminarily reflect the accuracy of relevant training, which is an important index in machine learning. It is affected by samples, noise intensity, privacy parameters and other factors. It can be inferred from this that the training intensity and model accuracy are jointly affected by noise, parameters and other factors.

3.2. Parameter-level sensitivity scoring mechanism

In the process of federated learning, clients usually perform several rounds of updates on the global model on local data, and upload the model parameter update amount to the server. Different parameters have significant differences in their change behaviors and gradient characteristics during training, and their impacts on privacy leakage risk and model performance are also inconsistent. Based on this situation, to realize refined differential privacy protection, it is necessary to describe the sensitivity of model updates at the parameter level. This study holds that a single index is difficult to fully reflect the importance and privacy risk of parameters, so it is necessary to incorporate both gradient information and parameter variation into the composition of sensitivity scoring.

3.2.1. Introduction of parameter variation characterization

Let the global model parameter issued by the server in the t-th round of federated training be w, which can be expressed as a whole:

$$w^{(t)} = \left\{ w^{(t)}{}_1, w_2^{(t)}, w_k^{(t)} \right\} \tag{3}$$

After the client completes training on the local dataset $\mathscr{D}_k$, the updated model $\mathbf{w}_k^{(t)}$ is obtained. The parameter update amount uploaded by the client is defined as the difference between the same parameters before and after one training:

$$w^{(t)} = w^{(t)} - w^{(t-1)} \tag{4}$$

Because the amplitude of parameter update can reflect the impact of local data on the current model. If the parameter changes greatly, it generally indicates that the parameter has a strong fitting ability to local samples, and it also means that it is easier to be used by attackers to restore the characteristics of training data and then obtain the original data. From this point of view, we can sort out the first part of the scoring mechanism. This paper first uses the absolute value of parameter variation to describe its potential sensitivity:

$$U_i^{[t]} = \left| \Delta W_{k\dot{i}}^{(t)} \right| \tag{5}$$

Among them, $\Delta w_{k,i}^{(t)}$ represents the update amount of the i-th parameter in the t-th round.

3.2.2. Introduction of gradient information characterization

It is difficult to accurately reflect the importance of this parameter in the optimization process only by the variation amplitude of parameters. Gradient information can reflect the constraint degree of

DOI: 10.54254/2753-8818/2026.31869

parameters on the current loss function. The larger the gradient amplitude is, the more important the parameter is in the current optimization direction. In this regard, if it is leaked, the possible privacy risk is also higher. In view of this, this paper uses the gradient of parameters calculated during local training:

$$G_{i,k}^{(t)} = \frac{\partial \mathcal{L}(w, D_k)}{\partial w_i} \tag{6}$$

Among them, $g_{k,i}^{(t)}$ represents the gradient tensor of a certain parameter in the t-th training iteration. ∂L represents the loss function.

3.2.3. Sensitivity scoring mechanism

After carefully considering both the behavioral performance of parameter changes and the importance of gradients, this paper combines the two for modeling, and defines the parameter-level sensitivity score as:

$$\mathrm{s}_{\mathrm{i}}^{(\mathrm{t})} = \mathrm{G}_{\mathrm{i}}^{(\mathrm{t})} \cdot \mathrm{U}_{\mathrm{i}}^{(\mathrm{t})} \tag{7}$$

If a parameter has both obvious changes in the current round of calculation and plays a certain role in the process of adjusting the loss function, its corresponding sensitivity score will increase significantly, and thus it will be identified as a high privacy risk parameter.

3.3. Mask generation

In the framework of selective differential privacy federated learning, the gradient-increment mask plays a role in connecting "sensitivity quantification" and "perturbation injection corresponding to differential privacy". In this regard, the main purpose of the mask is not to hide the parameters themselves, but to determine where to apply the corresponding perturbation and noise addition operation within the limited privacy budget, so as to avoid the performance loss caused by interfering with parameters without screening during all training processes.

3.3.1. Motivation for mask design

The traditional differential privacy federated learning method mostly adds uniform noise to each parameter update uploaded by clients. Although this global noise injection method can provide rigorous privacy guarantee, it will seriously disrupt the optimization direction of the model. In this regard, this problem is more obvious in the non-independent and identically distributed data scenarios.

In contrast, in the actual model training process, different parameters have obvious differences in their performance after model adjustment and the risk of privacy leakage. In the current round of training, only a small part of parameters have both strong gradient response and obvious update amplitude. In this regard, such parameters are more likely to be used by attackers for membership inference or attribute inference. Therefore, it is necessary to use a mechanism to screen parameters, so that the privacy budget can be concentrated on the most risky positions, so as to achieve a good protection effect.

Based on the above considerations, this paper starts to introduce the gradient-increment mask mechanism to identify the subset of parameters that need to be added with differential privacy

DOI: 10.54254/2753-8818/2026.31869

perturbation.

3.3.2. Mask generation based on sensitivity scoring

In the previous sensitivity quantification stage, the sensitivity score $s_i^{(t)}$ corresponding to each parameter position i has been calculated by the corresponding algorithm. The generation of the mask is to perform sorting and screening operations according to this corresponding sensitivity score.

Let the model parameter dimension be d. This study adopts the Top-k strategy to generate the mask vector $\mathbf{m}^{(t)} \in \{0,1\}^d$:

$$m_i^t = \begin{cases} 1, \ if \ s_i^t \in \text{Top-k} \\ 0, \ otherwise \end{cases} \tag{8}$$

Among them, $m_i^{(t)} = 1$ indicates that the parameter update position is selected for differential privacy noise addition, and k controls the sparsity of the mask. This design makes the mask scale adjustable, thus realizing flexible trade-off between privacy protection intensity, model accuracy and communication overhead.

This paper draws on the dynamic mask update method. After each round of local training, the sensitivity score will be recalculated and a new mask will be generated. In this regard, this dynamic method can make the mask keep up with the current sensitive area of the model, and make the privacy noise continuously act on the parameter updates that need protection most. In this regard, it can also avoid the privacy budget being accumulated on the already stable parameter positions for a long time.

The gradient-increment mask mentioned in this paper is completely generated locally, without mask alignment between different clients, and without server coordination. This design can reduce the system complexity, make privacy protection easier to achieve, and save the additional communication overhead caused by mask synchronization, reduce the transmission cost, and is more in line with the "local protection" design principle of differential privacy.

3.4. Dynamic differential privacy method

After completing the work related to parameter sensitivity scoring and mask generation, we then proceed to introduce the differential privacy mechanism to selectively add perturbation noise to the model update content uploaded by clients. In this regard, while doing a good job in privacy protection, we should minimize the drag on model performance. Different from the traditional method of adding noise to all parameters in federated learning, in this regard, the new architecture designed in this paper adopts a sensitivity-guided differential privacy injection strategy: for the update content corresponding to parameters with high sensitivity, noise perturbation is added. In this regard, other parameters maintain the original update form, so as to reduce the adverse effects caused by perturbation.

3.4.1. Noise addition objects and privacy protection goals

In federated learning, clients usually do not directly upload complete model parameters, but return the variation of the same parameter before and after one round of local training. Let the model

DOI: 10.54254/2753-8818/2026.31869

parameter of the client before the t-th round of training be $\mathbf{w}^{(t1)}$, and the parameter after training be $\mathbf{w}^{(t)}$. Then its uploaded update is defined as:

$$\boldsymbol{w}^{(\boldsymbol{t})} = \boldsymbol{w}^{(\boldsymbol{t})} - \boldsymbol{w}^{(\boldsymbol{t}-1)} \tag{9}$$

The differential privacy mechanism acts on the parameter update $\Delta\mathbf{w}^{(t)}$. By applying random noise to it, it is prevented that attackers infer the specific information of the client's local data from single or multiple model updates.

3.4.2. Selective differential privacy noise injection

Based on the gradient-increment mask vector $\mathbf{m}^{(t)} \in \{0,1\}^d$, differential privacy noise is only added to the parameter update positions marked as highly sensitive. Taking the Gaussian mechanism as an example, for the i-th parameter update, its perturbation injection process is defined as follows:

$$\Delta w_i^t = \begin{cases} \Delta w_i^t + N\left(0, \sigma_i^2\right), & m_i^t = 1 \\ \Delta w_i^t, & m_i^t = 0 \end{cases} \tag{10}$$

We try to add noise only to the update operations of high-sensitivity parameters, and then combine with the dynamic privacy budget allocation method. In terms of keeping the total privacy budget unchanged, we can significantly reduce the interference of the perturbed part on the overall optimization direction of the model. In this regard, theoretically, this strategy can reduce the addition of useless noise; in practical application, it also improves the convergence speed and final accuracy of the model.

4. Experimental results and analysis

4.1. Experimental environment and datasets

The experimental platform of this study is Windows 11 system, equipped with 32 GB memory and NVIDIA RTX5060 GPU. The experiment uses an environment based on PyTorch 2.2.1, and the dataset is MNIST. After 100 rounds of training, the digital pictures in the data are detected to judge the model accuracy.

4.2. Model accuracy comparison experiment

To verify the effectiveness of the sensitivity-guided selective differential privacy federated learning method proposed in this paper, this study conducts simulation experiments under a classic federated learning experimental framework. The experiment adopts a multi-client non-independent and identically distributed data division method to simulate the real federated learning scenario. In each round of training, some clients are randomly selected to participate in model update, and the server aggregates parameters relying on the federated averaging algorithm. The experimental results are shown in Table 1 and Fig. 2.

DOI: 10.54254/2753-8818/2026.31869

Table 1. Comparison of model accuracy under different privacy mechanisms

Method	Test Accuracy (%)
Non-DP FL	92.4
Global-DP FL	84.7
Proposed Method	90.8

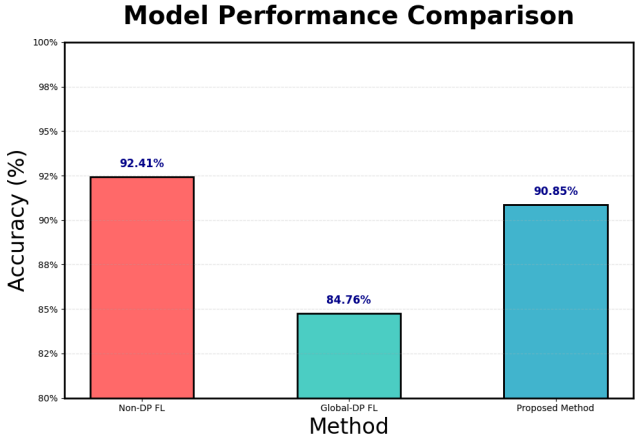

Figure 2. Comparison chart of model accuracy

It can be seen from the experimental results that the traditional full-parameter differential privacy method introduces a large amount of indiscriminate noise under the setting of the same privacy budget, and equally disturbs all parameters, which significantly limits the model optimization process and leads to a significant decline in model accuracy. In this regard, the method in this paper applies perturbation only to the update operations of high-risk parameters with the help of sensitivity scoring and mask mechanism, and significantly reduces the loss of accuracy on the premise of ensuring privacy protection.

In this regard, the accuracy of the constructed method is relatively close to that of federated learning without privacy protection, which shows that the selective perturbation injection operation can effectively retain the main discriminative ability of the model.

4.3. Communication overhead comparison experiment

Due to the introduction of the mask mechanism, the method in this paper only injects noise into part of the parameter updates, and allows direct upload or compressed transmission of low-sensitivity parameters. Table 2 shows the average proportion of uploaded parameters in single-round communication of different methods.

Table 2. Comparison of communication overhead of different methods

Method	Proportion of Uploaded Parameters
Global-DP FL	100%
Proposed Method	35%

Experimental results show that the mask mechanism can significantly reduce communication overhead while ensuring the privacy protection effect, which is especially suitable for edge devices

DOI: 10.54254/2753-8818/2026.31869

with limited bandwidth and cross-institutional collaboration scenarios.

4.4. Result discussion

Comprehensive experimental results can draw the following conclusions:

1. Under the constraint of the same privacy budget, the selective differential privacy method is significantly superior to the traditional global noise addition scheme in model accuracy;
2. The mask mechanism can effectively reduce communication overhead and improve the deployment feasibility of federated learning system;
3. The sensitivity-guided noise scheduling strategy makes privacy protection more targeted and avoids invalid perturbation injection.

The above results verify that the method in this paper has achieved a good balance among privacy protection, model performance and system efficiency.

5. Conclusion

Aiming at the problem that it is difficult to achieve both privacy protection and model performance in federated learning, this paper constructs a selective differential privacy method guided by sensitivity. We model parameter gradients and update amplitudes together to establish a parameter-level sensitivity scoring mechanism, and then generate gradient-increment masks according to this mechanism to accurately add differential privacy noise to the positions of high-risk parameters. This method significantly reduces the noise interference on irrelevant parameters under the condition of complying with the theoretical constraints of differential privacy, thus significantly alleviating the performance degradation problem caused by the traditional global noise addition method.

Experimental results show that under the same privacy budget setting, the method proposed in this paper performs better than the compared schemes in terms of model accuracy, convergence speed and communication efficiency, which also shows that it has certain effectiveness and practicability in non-independent and identically distributed federated scenarios. The future research of ours will carefully study more refined privacy budget scheduling strategies and theoretical analysis related to adversarial inference attacks, so as to enhance the applicability and robustness of the above strategies in large-scale federated learning systems running in real scenarios.

References

[1] Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In Theory of Cryptography Conference (TCC) (pp. 265-284). Springer, Berlin, Heidelberg.

[2] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (pp. 308-318).

[3] Wu, X., Li, F., Kumar, A., Chaudhuri, K., Jha, S., & Naughton, J. (2017). Bolt-on differential privacy for scalable stochastic gradient descent-based analytics. In Proceedings of the 2017 ACM International Conference on Management of Data (pp. 1307-1322).

[4] Liu, X., Li, Y., & Wang, S. (2018). Adaptive differential privacy for distributed machine learning. IEEE Transactions on Information Forensics and Security, 14(2), 460-474.

[5] Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., ... & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. IEEE Transactions on Information Forensics and Security, 15, 3454-3469.

[6] Zhang, T., Zhu, Q., & Yu, S. (2019). Sparse differential privacy for high-dimensional data publishing. IEEE Transactions on Knowledge and Data Engineering, 32(10), 2006-2019.

DOI: 10.54254/2753-8818/2026.31869

[7] Jiang, Y., Zhou, Z., & Chen, Q. (2021). Limitations of random masking in differential privacy: A theoretical perspective. Advances in Neural Information Processing Systems, 34, 12345-12358.
[8] Wang, Z., Huang, Y., & Liu, Y. (2022). Gradient-aware selective differential privacy for federated learning. In International Conference on Machine Learning (pp. 22765-22778). PMLR.
[9] Chen, R., Zhang, S., & Li, D. (2023). Gradient instability in non-convex differential privacy: Challenges and mitigations. In International Conference on Learning Representations.