Scientific Research Publishing

Journal of Information Security, 2023, 14, 111-135
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

A Systematic Survey for Differential Privacy Techniques in Federated Learning

Yi Zhang[1,2], Yunfan Lu[2,3*], Fengxia Liu[4,5*]

[1]Institute for Mathematical Sciences, Renmin University of China, Beijing, China
[2]Engineering Research Center of Financial Computing and Digital Engineering, Ministry of Education, Beijing, China
[3]School of Mathematics, Renmin University of China, Beijing, China
[4]Institute of Artificial Intelligence, Beihang University, Beijing, China
[5]Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing, Beijing, China
Email: *luyunfan@ruc.edu.cn, *shunliliu@buaa.edu.cn

How to cite this paper: Zhang, Y., Lu, Y.F. and Liu, F.X. (2023) A Systematic Survey for Differential Privacy Techniques in Federated Learning. *Journal of Information Security*, **14**, 111-135.
https://doi.org/10.4236/jis.2023.142008

Received: December 13, 2022
Accepted: February 25, 2023
Published: February 28, 2023

Abstract

Federated learning is a distributed machine learning technique that trains a global model by exchanging model parameters or intermediate results among multiple data sources. Although federated learning achieves physical isolation of data, the local data of federated learning clients are still at risk of leakage under the attack of malicious individuals. For this reason, combining data protection techniques (e.g., differential privacy techniques) with federated learning is a sure way to further improve the data security of federated learning models. In this survey, we review recent advances in the research of differentially-private federated learning models. First, we introduce the workflow of federated learning and the theoretical basis of differential privacy. Then, we review three differentially-private federated learning paradigms: central differential privacy, local differential privacy, and distributed differential privacy. After this, we review the algorithmic optimization and communication cost optimization of federated learning models with differential privacy. Finally, we review the applications of federated learning models with differential privacy in various domains. By systematically summarizing the existing research, we propose future research opportunities.

Keywords

Federated Learning, Differential Privacy, Privacy Computing

1. Introduction

Machine learning is a branch of artificial intelligence and computer science that focuses on using data and algorithms to mimic how humans learn (Bishop and

Nasrabadi [1]). Many fields (e.g., natural language processing (Nadkarni *et al.* [2]), computer vision (Jarvis [3]), bioinformatics (Fatima and Pasha [4]), etc.) have benefited from machine learning. An imperative prerequisite to the success of machine learning is the availability of large amounts of high-quality data. With the help of the data, machine learning models can discover patterns in the data and perform tasks that are difficult for humans to carry out, such as fraud detection (Bolton and Hand [5]), face recognition (Zhao *et al.* [6]), speech recognition (Reddy [7]), etc.

Machine learning models are centralized models, meaning that all the data for the training of the model must be centralized in one location (e.g., a data center). As a result, machine learning models face a number of challenges in practice. The first challenge is the breach of privacy (Ji *et al.* [8]). Data privacy breaches have been observed in numerous examples as a result of centralizing personal data in one place (Ouadrhiri and Abdelhadi [9]). Google+, for instance, has leaked personally identifiable information such as name, email address, occupation, gender, age, and relationship status of approximately 500,000 users because of a system vulnerability in 2018. Facebook compromised the personal information of approximately 533 million users in 2021 (e.g., phone numbers, login IDs, full names, etc.). Twitter compromised 5.4 million accounts in 2022, including phone numbers, locations, URLs, profile pictures, etc. The second challenge is the problem of data silos. Several countries around the world have passed strict laws to safeguard the privacy of personal information. These laws include the General Data Protection Regulation of the European Union and the Data Security Law of the People's Republic of China. While these laws protect the security of data, they also restrict the flow of data, resulting in the problem of data silos.

A major challenge in the field of artificial intelligence today is how to resolve the problem of data silos while maintaining data privacy and security. Towards this end, Google proposed federation learning in 2017 (McMahan *et al.* [10]). The federated learning model is a distributed machine learning model that enables multiple devices to collaboratively train the same machine learning model without exchanging local data, only by exchanging model parameters or intermediate results, thus achieving a balance between data privacy protection and data sharing computation. In the federated learning model, the data of each participant is stored locally instead of being centralized in the central server, so that the security of each participant's data can be maintained to a certain extent. After the introduction of federated learning, it has rapidly received attention from both academia and industry, and many mature federated learning frameworks have been developed, for example, TensorFlow Federated developed by Google, Pysyft developed by OpenMinded (Ziller *et al.* [11]), and FATE developed by Webank (Liu *et al.* [12]).

While the data of each participant is stored locally in a federated learning framework, there is still a risk of data leakage (Li *et al.* [13]). In the course of training a federated learning model, each participant must transmit information about the model parameters (e.g., gradients) to the central server. A number of

examples show that the central server can invert participants' local data using gradient information, which is known as an inference attack (Nasr *et al.* [14]). To make the federated learning model more secure, scholars have introduced differential privacy techniques into federated learning. As a result of the combination of differential privacy techniques with federation learning, the data leakage problem of federation learning models can be effectively solved.

Although there are some literature reviews on differential-private federated learning models, there are some shortcomings. First, federed learning is developing rapidly and many of the latest research results have not been reviewed in, so this paper summarizes the latest research results on differential-private federated learning. Second, the research on the optimization of differential-private federated learning models is also important for the development of this field, but less attention has been paid by scholars. Therefore, this paper summarizes the research related to the optimization techniques for differential-private federated learning models. Finally, differential-private federated learning has achieved applications in many fields, which this paper summarizes.

This paper reviews the recent advances in differential privacy techniques for federated learning. The rest of this paper is structured as follows. In Section 2, we review some background knowledge about federated learning and differential privacy. In Section 3, we review the recent advances in federated learning with central differential privacy, local differential privacy, and distributed differential privacy, respectively. In Section 4, we review the algorithm optimization techniques and communication cost optimization techniques in the differential private federated learning model. In Section 5, we review some recent applications of the differential private federated learning model. In Section 6, we propose some future directions. In Section 7, we draw conclusions.

2. Fundamental Principles

In this section, we review some of the basic concepts of federated learning and differential privacy.

2.1. Federated Learning

In a federated learning model, multiple devices or clients collaborate in training a machine learning model by exchanging only model parameters or intermediate results without exchanging local data. As a result, federated learning consists of two key components: the central server and the federated learning clients. Let $\mathcal{N}=\{1,2,\cdots,N\}$ be the set of federated learning clients. Each client has a local data set $\mathcal{D}_i$ with data structure $\left(\boldsymbol{x}_j,\boldsymbol{y}_j\right)$, where $\boldsymbol{x}_j\in\mathbb{R}^{d_1}$ represents the features and $\boldsymbol{y}_j\in\mathbb{R}^m$ represents the labels. The central server is responsible to decide the architecture of the machine learning model (e.g., logistic regression, etc.), and then sends the model information and model initialization parameters $\boldsymbol{\omega}_0\in\mathbb{R}^{d_2}$ to all federated learning clients. According to the central server, each client downloads the initial parameters, trains the machine learning model with

their local data, and uploads the model parameters to the central server after the training is completed. The central server aggregates the parameters uploaded by all clients to form the global model parameters. In order to make the trained model more effective, the above process is performed several times until the model converges. Specifically, the federation learning model can be divided into the following three steps: *initialization*, *local training*, and *global aggregation*, and the workflow of the federated learning is shown in Figure 1.

Step 1. (Initialization) The central server decides the architecture of the machine learning model and sends the initial parameters $\boldsymbol{\omega}_0$ to each client.

Step 2. (Local Training) In the t-th round, client i downloads the parameter $\boldsymbol{\omega}_{t-1}$ from the central server, and updates the parameter by minimizing its loss function based on its local data set $\mathcal{D}_i$, *i.e.*,

$$\boldsymbol{\omega}_t^{i*} = \arg\min_{\boldsymbol{\omega}_{t-1}^i} \left(L_i \left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1} \right) \right), \tag{1}$$

$$L_i \left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1} \right) = \sum_{j=1}^{|\mathcal{D}_i|} l \left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1} \right), \tag{2}$$

where $l : \mathbb{R}^{d_1} \times \mathbb{R}^m \times \mathbb{R}^{d_2} \to \mathbb{R}$ is the loss function for data sample j and is dependent on the underlying machine learning model; $L_i : \mathbb{R}^{d_1} \times \mathbb{R}^m \times \mathbb{R}^{d_2} \to \mathbb{R}$ is the loss function for client i; $|\mathcal{D}_i|$ is the number of the data in $\mathcal{D}_i$. Equations (1) and (2) are usually solved by the stochastic gradient descent method, *i.e.*,

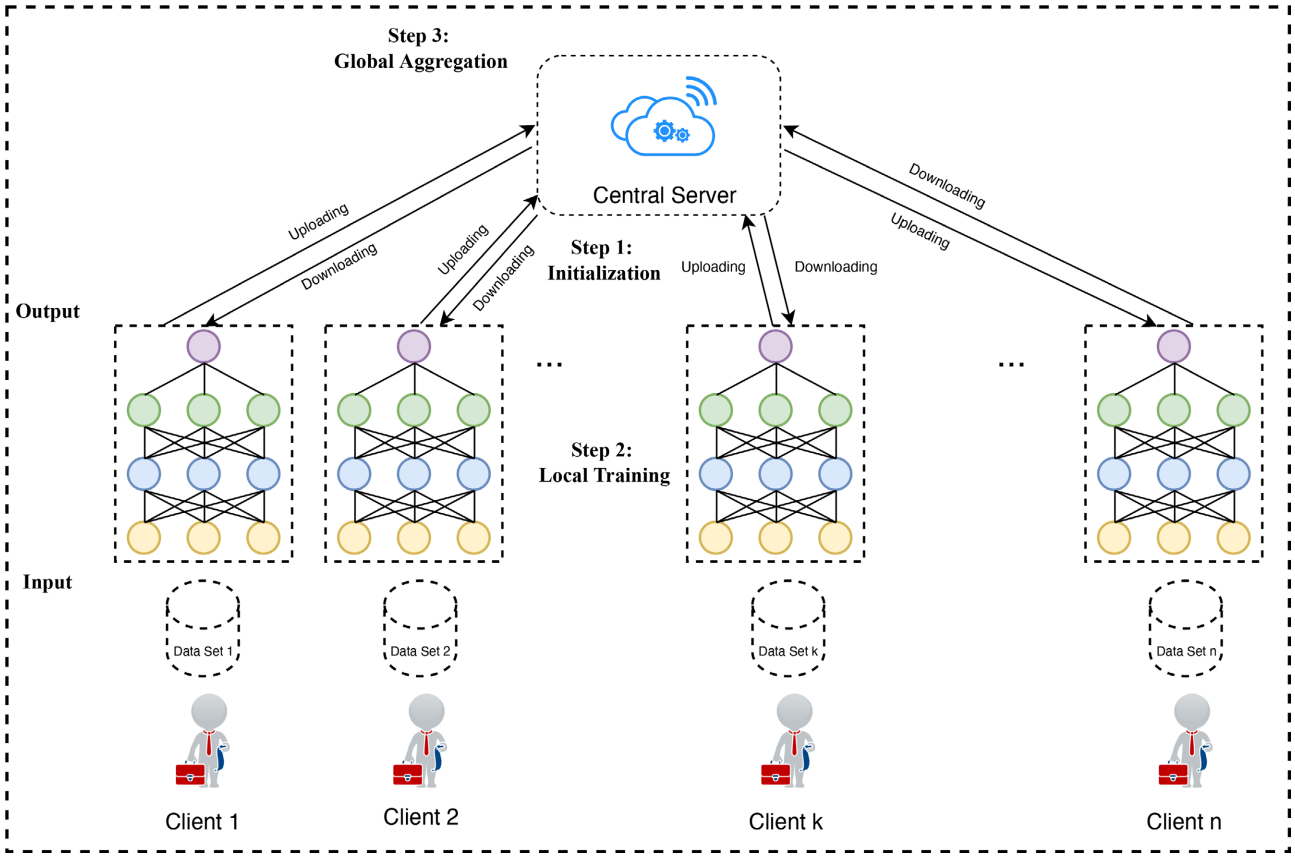

Figure 1. Workflow of federated learning.

$$\boldsymbol{\omega}_t^{i*} = \boldsymbol{\omega}_{t-1}^{i*} - \eta \nabla L_i\left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{w}_{t-1}\right) \tag{3}$$

where $\nabla L_i\left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1}\right)$ is the gradient of $L_i\left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1}\right)$ and η is the learning rate. Client i uploads the intermediate results (e.g., $\nabla L_i\left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1}\right)$) to the central server.

Step 3. (Global Aggregation) The central server collects the intermediate results for each client and updates the global model parameters $\boldsymbol{\omega}_t$ through the global model aggregation algorithm. For example, under the FedAVG [10],

$$\boldsymbol{\omega}_t^* = \boldsymbol{\omega}_{t-1}^* - \eta \sum_i^N \frac{\left|\mathcal{D}_i\right|}{\sum_i^N \left|\mathcal{D}_i\right|} \nabla L_i\left(\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{\omega}_{t-1}\right). \tag{4}$$

In addition to FedAVG, many variants of FedAVG (e.g., FedProx [15], FedPAQ [16], Turbo-Aggregate [17], FedMA [18], HierFAVG [19]) can also be used.

For more details on federated learning models, the readers may refer to Yang *et al.* [20], Rehman and Gaber [21], and Ludwig and Baracaldo [22].

2.2. Differential Privacy

Differential privacy (Dwork and Roth [23]) is a data protection technique based on probability theory, and the idea behind it is that if for two adjacent databases (*i.e.*, two databases differing by only one record), the statistical characteristics derived from these two databases cannot be used to deduce the single record, then the records in this database are said to be secure.

To this end, Dwork *et al.* [23] first give the definition of distance between databases and the definition of the randomized algorithm, and then give the concept of differential privacy. A randomized algorithm is an algorithm with the domain A and (discrete) range B will be associated with a mapping from A to the probability simplex over B, denoted $\Delta(B)$, and $\Delta(B) = \left\{x \in \mathbb{R}^{|B|} : x_i \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^{|B|} x_i = 1\right\}$. For the databases x and y being collections of records from a universe $\mathcal{X}$ and being represented by their histograms (*i.e.*, $x \in \mathbb{N}^{|\mathcal{X}|}$, in which each entry x_i represents the number of elements in the database x of type $i \in \mathcal{X}$), and the distance between x and y can be given by $\|x-y\|_1$, where $\|\cdot\|_1$ is the l_1-norm.

Definition 1. (Dwork *et al.* [23]) A randomized algorithm $\mathcal{M}$ with domain $\mathbb{N}^{|\mathcal{X}|}$ is (ε, δ)-differentially private if for all $\mathcal{S} \subseteq \text{Range}(\mathcal{M})$ and for all $x, y \in \mathbb{N}^{|\mathcal{X}|}$ such that $\|x-y\|_1 \leq 1$:

$$\Pr\left[\mathcal{M}(x) \in \mathcal{S}\right] \leq \exp(\varepsilon) \Pr\left[\mathcal{M}(y) \in \mathcal{S}\right] + \delta. \tag{5}$$

As described above, the definition of differential privacy guarantees privacy theoretically, but implementation requires perturbing the data by adding noise. By defining

$$\Delta f = \max_{\substack{x, y \in \mathbb{N}^{|\mathcal{X}|} \\ \|x-y\|_1 = 1}} \left\| f(x) - f(y) \right\|_1 \tag{6}$$

as the ℓ_1-sensitivity of a deterministic algorithm $f:\mathbb{N}^{|\mathcal{X}|}\to\mathbb{R}^k$, Dwork and Roth [23] add a random variable $Y\sim \mathrm{Lap}(\Delta f/\varepsilon)$ to the deterministic algorithm and propose a Laplacian mechanism, and prove that the Laplace mechanism preserves $(\varepsilon,0)$-differential privacy.

Definition 2. (Dwork *et al.* [23]). Given a deterministic algorithm $f:\mathbb{N}^{|\mathcal{X}|}\to\mathbb{R}^k$, the Laplace mechanism is defined as:

$$\mathcal{M}_L\left(x,f(\cdot),\varepsilon\right)=f(x)+\left(Y_1,\cdots,Y_k\right) \tag{7}$$

where Y_i are i.i.d. random variables drawn from $\mathrm{Lap}(\Delta f/\varepsilon)$.

Theorem 1 (Dwork *et al.* [23]). The Laplace mechanism preserves $(\varepsilon,0)$-differential privacy.

By defining the arbitrary sensitivity

$$\Delta u\equiv\max_{r\in\mathcal{R}}\max_{x,y:\|x-y\|_1\le 1}\left|u(x,r)-u(y,r)\right|, \tag{8}$$

McSherry and Talwar [24] propose the exponential mechanism and proved that the exponential mechanism preserves $(\varepsilon,0)$-differential privacy.

Definition 3. (McSherry and Talwar [24]). The exponential mechanism $\mathcal{M}_E(x,u,\mathcal{R})$ selects and outputs an element $r\in\mathcal{R}$ with probability proportional to $\exp\left(\frac{\varepsilon u(x,r)}{2\Delta u}\right)$.

Theorem 2. (McSherry and Talwar [24]). The exponential mechanism preserves $(\varepsilon,0)$-differential privacy.

By defining

$$\Delta_2 f=\max_{\substack{x,y\in\mathbb{N}^{|\mathcal{X}|}\\ \|x-y\|_1=1}}\left\|f(x)-f(y)\right\|_2 \tag{9}$$

as the l_2-sensitivity of a deterministic algorithm $f:\mathbb{N}^{|\mathcal{X}|}\to\mathbb{R}^k$, Nikolov *et al.* [25] add a random variable $Y\sim N\left(0,\sigma^2\right)$ and $\sigma\ge\sqrt{2\ln\frac{1.25}{\delta}}\cdot\frac{\Delta_2 f}{\varepsilon}$ to the deterministic algorithm and propose a Gaussian mechanism, and prove that the Gaussian mechanism preserves (ε,δ)-differential privacy.

Definition 4. (Nikolov *et al.* [25]). Given a deterministic algorithm $f:\mathbb{N}^{|\mathcal{X}|}\to\mathbb{R}^k$, the Gaussian mechanism is defined as:

$$\mathcal{M}_L\left(x,f(\cdot),\varepsilon\right)=f(x)+\left(Y_1,\cdots,Y_k\right) \tag{10}$$

where Y_i are i.i.d. random variables drawn from $N\left(0,\sigma^2\right)$ and $\sigma\ge\sqrt{2\ln\frac{1.25}{\delta}}\cdot\frac{\Delta_2 f}{\varepsilon}$.

Theorem 3. (Nikolov *et al.* [25]). Let $\varepsilon\in(0,1)$ be arbitrary. For $c^2>2\ln(1.25/\delta)$, the Gaussian Mechanism with parameter $\sigma\ge c\Delta_2 f/\varepsilon$ is (ε,δ)-differentially private.

For more details about differential privacy, the readers may refer to Dwork [26], Dwork and Roth [27], and Ji *et al.* [8].

3. An Overview of Federated Learning with Differential Privacy

During the federated learning process, the federated clients need to transmit parameters (e.g., gradients) to the central server, which may lead to the leakage of the federated clients' local data. In order to protect the data of federated clients, both federated clients and the central server must use data protection techniques. Differential privacy is a probabilistic-based data privacy protection technique that has been successful in federated learning. The studies of differential privacy techniques in federated learning can be divided into three categories: federated learning with central differential privacy, federated learning with local differential privacy, and federated learning with distributed differential privacy. In this section, we provide an overview of federation learning with differential privacy. Specifically, we will first review the recent advances in federated learning with central differential privacy (in Subsection 3.1). Then, we will review recent advances in federation learning with local differential privacy (in Subsection 3.2). Finally, we will review recent advances in federated learning with distributed differential privacy (in Subsection 3.3).

3.1. Federeated Learning with Central Differential Privacy

Federated learning with central differential privacy is the way that a trusted central server adds noise to global parameters to protect local data. The workflow federated learning with central differential privacy is shown in **Figure 2**.

Geyer *et al.* [28] note that vanilla federated learning can be subject to differential attacks, thus initiating the study of federated learning with central differential privacy. In particular, a trusted server adds noise to aggregate results in order to protect against differential attacks. According to numerical experiments, this approach provides data security protection at the expense of accuracy. Triastcyn and Faltings [29] propose Bayesian differential privacy as a means of providing more precise privacy loss bounds. As demonstrated in experiments, the bayesian differential privacy significantly reduces noise, improves model accuracy, and reduces the number of communication rounds. The proposed method improves the accuracy of trained models by up to 10% according to experimental results. In Wei *et al.* [30], they propose a novel approach, NbAFL, which adds artificial noise to parameters at the client's side prior to aggregation. The NbAFL can satisfy the central DP under different levels of protection by properly adapting different variances of artificial noise. Furthermore, the authors develop a theoretical convergence bound for the loss function of the trained FL model in the NbAFL. Bernau *et al.* [31] examine the inference attack on central differential privacy. The authors of Zhang *et al.* [32] propose a clipping-enabled FedAvg, which combines the clipping technique with federated learning and central differential privacy. They demonstrate the relationship between clipping bias and the distribution of the client's updates by analyzing the convergence of FedAvg with clipping. Hu *et al.* [33] present a new differentially-private FL scheme

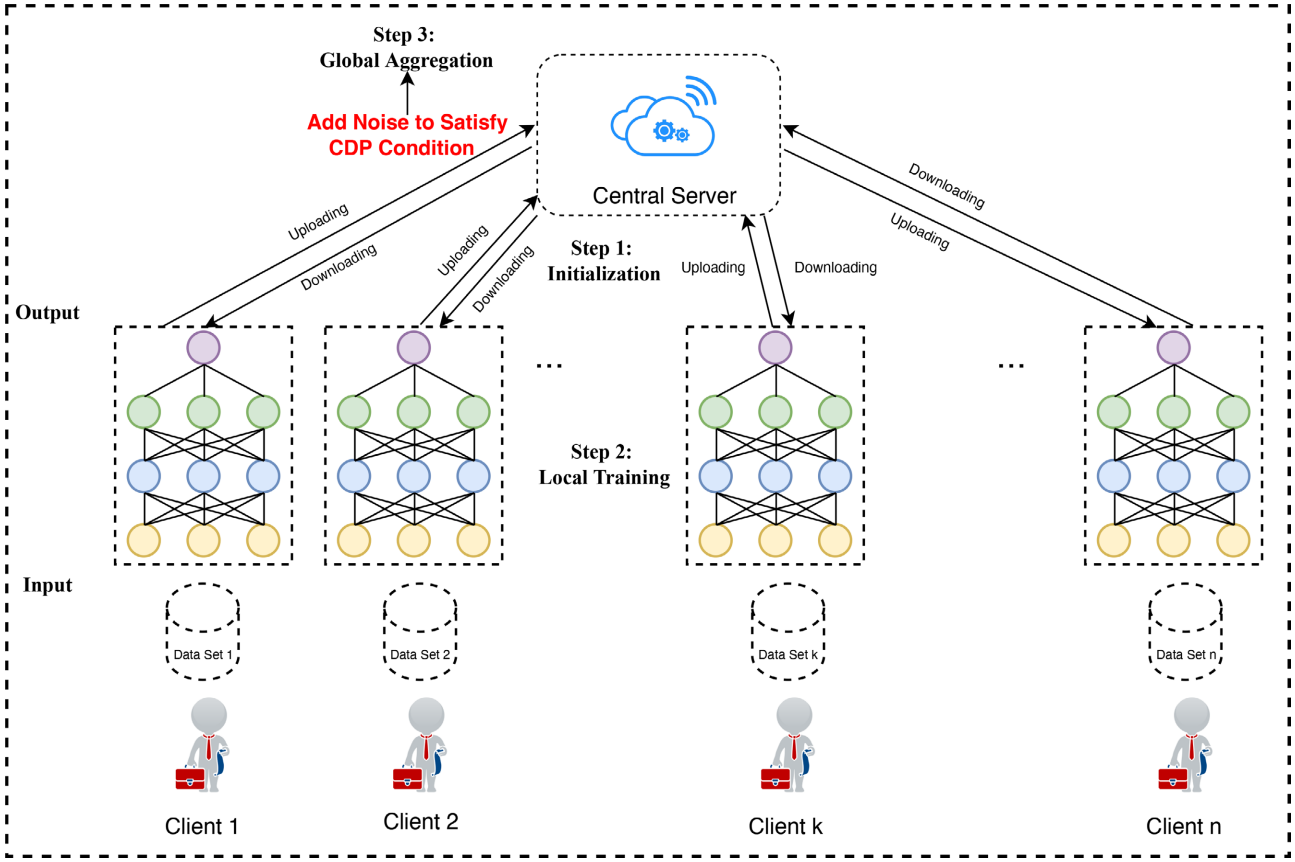

Figure 2. Workflow of federated learning with central differential privacy.

referred to as Fed-SMP, which provides client-level DP guarantees while maintaining high model accuracy. To minimize the impact of privacy protection on model accuracy, Federal-SMP employs a new technique called Sparsified Model Perturbation (SMP), which involves sparsifying local models before perturbing them with additive Gaussian noise. Extensive experiments on real-world datasets demonstrate Fed-SMP's capability to improve model accuracy while simultaneously reducing communication costs. Table 1 summarizes the recent advance in central differential privacy.

3.2. Federated Learning with Local Differential Privacy

In the framework of federation learning with central differential privacy, a necessary condition for this framework to be able to secure client data is that the central server is trusted, and if the central server is honest but curious, then the local client's data will be leaked to the central server (e.g., Li *et al.* [34] and Melis *et al.* [35]). Therefore, a more secure framework is the federation learning with local differential privacy, *i.e.*, each client adds noise to the parameters uploaded to the central server to secure the local data. The workflow federated learning with local differential privacy is shown in Figure 3.

Federated learning with local differential privacy is first formalized by Kasiviswanathan *et al.* [36]. They show that a concept class is learnable by a local

Table 1. Summary of contributions in central differential privacy.

Approach	Technique	Main Idea
Geyer *et al.* [28]	FL + CDP	Initiate the study of federated learning with central differential privacy and verify the validity of the model by numerical experiments
Triastcyn and Faltings [29]	Bayesian differential privacy	Develop a relaxation of federated learning with central differential privacy, named Bayesian differential privacy.
Wei *et al.* [30]	NbAFL	Develop a novel framework (named NbAFL) based on DP in which artificial noises are added
Zhang *et al.* [32]	Clipping-enabled FedAvg	Develop a novel central differential privacy framework (named clipping-enabled FedAvg) based on clipping technique.
Hu *et al.* [33]	Fed-SMP	Develop a novel framework, named Fed-SMP, to mitigate the inaccuracy issue of LDP by using a technique called Sparsified Model Perturbation (SMP) where local models are sparsified first before being perturbed by Gaussian noise.

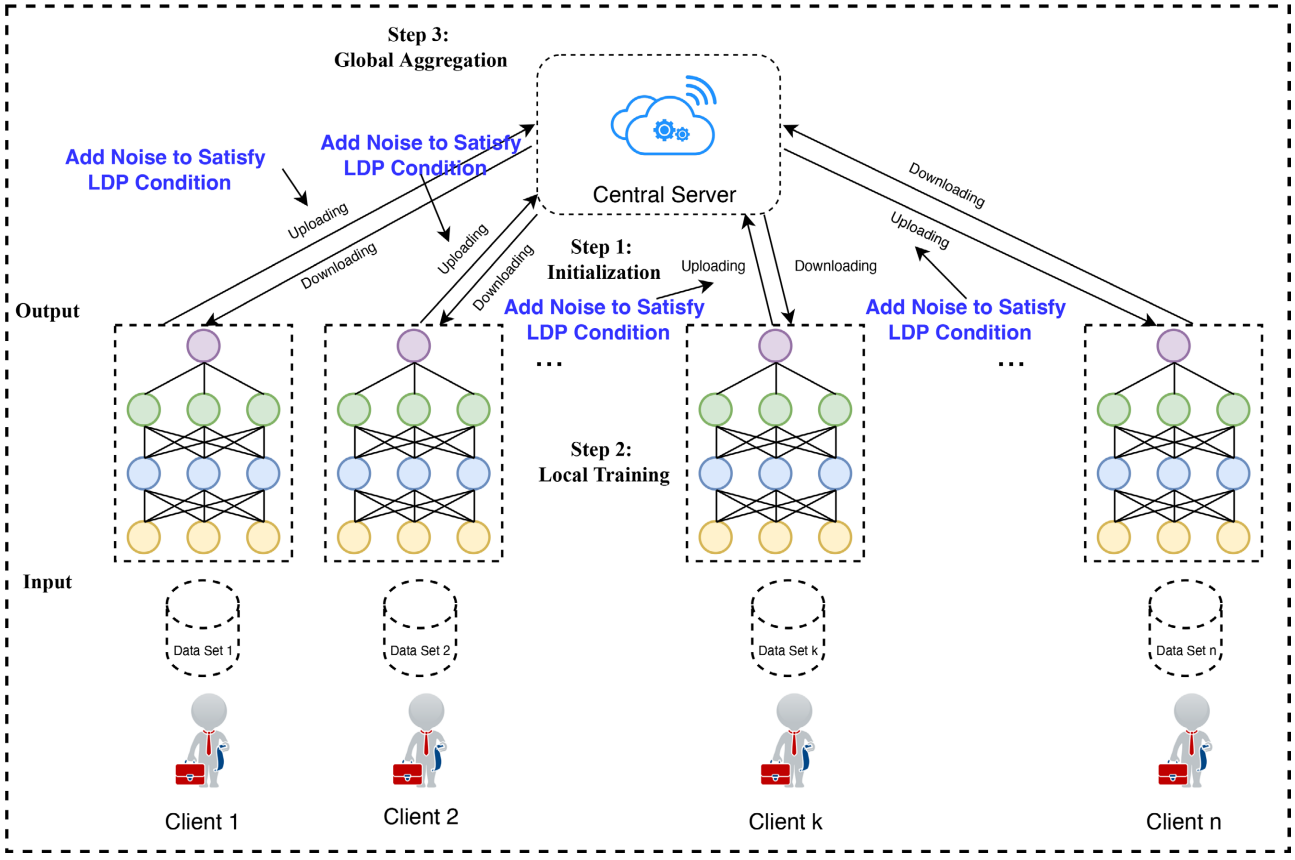

Figure 3. Workflow of federated learning with local differential privacy.

differentially private algorithm if and only if it is learnable in the statistical query model. Erlingsson *et al.* [37] propose a privacy-preserving mechanism called

Randomized Aggregatable Privacy-Preserving Ordinal Response, or RAPPOR. They demonstrate that RAPPOR allows the collection of statistics on the population of client-side strings with strong privacy guarantees for each client and without linking the reports of the clients. Liu [38] proposes a generalized Gaussian (GG) mechanism based on LP global sensitivity and demonstrates that the GG mechanism reaches DP at a specified level of privacy. Truex *et al.* [39] focuses on federated learning frameworks with high-dimensional, continuous values and high-precision client data. As a result, the existing LDP protocols cannot be applied in this situation. The authors therefore proposed LDP-Fed, which provides a formal differential privacy guarantee for repeated collection of model parameters in federated neural network training over multiple individual participants' private datasets. Sun *et al.* [40] examine whether differential privacy can protect backdoor attacks and demonstrate that norm clipping and weak differential privacy mitigate attacks without affecting overall performance. Xu *et al.* [41] addresses the situation where sensitive data about each user must be collected from multiple services independently and can be combined. In this research, the authors focus on preventing the privacy guarantee from being compromised during the joint collection of data and on how to analyze perturbed data from different services jointly. Towards this end, they propose mechanisms and estimation methods to process multidimensional analytical queries. Naseri *et al.* [42] investigate the robustness of local differential privacy techniques in FL. Experiments show that central differential privacy techniques are robust to defend against backdoor attacks. Wang *et al.* [43] propose a local differential privacy-based framework (named FedLDA) for federated learning of LDA models, as well as a novel LDP mechanism called Random Response with Priori (RRP). According to theoretical results, the novel framework provides theoretical guarantees regarding data privacy as well as model accuracy. Girgis *et al.* (2021) [44] propose a novel shuffle privacy model, in which each client randomizes its response and the server only receives a random shuffle of the clients' responses. In sub-sampled shuffled models, numerical results demonstrate significant improvements in privacy guarantee over the state-of-the-art approximate Differential Privacy guarantee. Wei *et al.* [45] proposes a user-level differential privacy (UDP) algorithm by adding artificial noise to the shared models before they are uploaded to the servers. Through varying the variances of the artificial noise processes, they demonstrate that the UDP framework can achieve (ε-δ)-LDP for the ith mobile terminal with adjustable privacy levels. Furthermore, they derive a theoretical upper bound for the convergence of UDP. Zhou *et al.* [46] propose a novel privacy-preserving federated learning framework for edge computing (PFLF). In PFLF, each client and the central server add noise before sending the data. For the purpose of protecting the privacy of their clients, they developed a flexible arrangement mechanism for counting the optimal training times for each individual, and prove that PFLF guarantees the privacy of clients and servers during the entire training process. As Thapa *et al.* [47] observe, federated

learning and split learning are two distributed machine learning methods that perform similar tasks. Federated learning does not provide as much privacy as split learning. Split learning, however, performs slower than Federated learning. Thapa *et al.* [47] combines federated learning with split learning and presents a novel approach, splitfed learning (SFL), as well as a revised architectural configuration that incorporates differential privacy to enhance data privacy and model robustness. Wu *et al.* [48] present FedPerGNN, a federated GNN framework, which is capable of both effective and privacy-preserving personalization. The experimental results on six datasets demonstrate that FedPerGNN is capable of achieving 4.0 - 9.6 percent lower errors than the state-of-the-art federated personalization methods under good privacy protection. To secure cross-silo federated learning, Wang *et al.* [49] proposes a three-plane approach in which Local Differential Privacy is applied to user data before it is uploaded. According to theoretical results, LDP is capable of providing strong data privacy protection and still retaining user data statistics in order to maintain its high utility. Zhang *et al.* [50] propose federated f-differential privacy, a new notion specifically tailored to federated settings, based on the framework of Gaussian differential privacy. They then design PriFedSync as a generic framework for private federated learning. **Table 2** summarizes the recent advance in local differential privacy.

Table 2. Summary of contributions in local differential privacy.

Approach	Technique	Main Idea
Wang *et al.* [43]	FedLDA	Develop a novel framework (named FedLDA) for federated learning of LDA models, as well as a novel LDP mechanism called Random Response with Priori (RRP).
Girgis *et al.* [44]	Shuffle method	Develop shuffle method under with only a random permutation of the clients' responses are received by the server without their association with the clients' identities.
Wei *et al.* [45]	UDP	Develop a novel framework, named UDP, in which each client can achieve adjustable privacy protection levels.
Zhou *et al.* [46]	PFLF	Develop a novel framework, named PFLF, in which the client and the central server add noise before sending the data.
Thapa *et al.* [47]	SFL	Develop a novel framework SFL that combines federated learning and split learning to achieve high model accuracy and communication efficiency.
Wu *et al.* [48]	FedPerGNN	Develop a novel framework FedPerGNN to achieve both effective and privacy-preserving personalization.
Zhang *et al.* [50]	PriFedSync	Develop federated f-differential privacy and propose a generic framework for private federated learning.

3.3. Federated Learning with Distributed Differential Privacy

Both centralized differential privacy and local differential privacy have shortcomings. For centralized differential privacy, it requires a trusted central server, and once the central server is malicious, then the data of the federated learning clients will be compromised, and a trusted central server is hard to find in practice. For local differential privacy, each client adds a lot of noise to the intermediate results they upload to satisfy the local differential privacy condition. Although the local data of the federated learning clients are secure under local differential privacy, it leads to too much noise in the aggregated results obtained by the federated servers, which leads to poor privacy-utility trade-offs. To address these shortcomings, scholars have proposed a differential privacy model that can guarantee data security while keeping the amount of added noise limited, *i.e.*, distributed differential privacy. In this model, each federated learning client only needs to add a small amount of noise to ensure that the aggregation result of the central server satisfies the central differential privacy; at the same time, since the amount of noise added by each federated learning client is small and cannot guarantee the security of local data, the federated clients will use the secure aggregation technique (Bonawitz *et al.* (2017) [51]), so that the federated server can only get the aggregation result of the intermediate parameters of all federated clients, but cannot get the intermediate parameters of each federated client, thus securing the federated clients' local data. The workflow federated learning with distributed differential privacy is shown in **Figure 4**.

Dwork *et al.* [52] proposed a Binomial mechanism and prove that the binomial mechanism achieves the $(\varepsilon - \delta)$ differential privacy condition. Agarwal *et al.* [53] applied the binomial mechanism to federated learning and proposed a stochastic k-level quantization method and a randomized rotation method. Results show that the Binomial mechanism with the stochastic k-level quantization method and randomized rotation method can achieve nearly the same utility as the Gaussian mechanism, yet requires fewer representation bits. Canonne *et al.* [54] proposed a discrete Gaussian mechanism. According to the authors, discrete Gaussian noise can provide essentially the same level of privacy and accuracy as continuous Gaussian noise, both theoretically and experimentally. On the basis of Canonne *et al.* [54]'s work, Kairouz *et al.* [55] applied the discrete Gaussian mechanism to federated learning with secure aggregation, and results show that the model can provide $1/2\varepsilon^2$ central differential privacy and the mean squared error of the model is at most $O\left(c^2d^2/\varepsilon^2\right)$. Agarwal *et al.* [56] proposes a new multi-dimensional Skellam mechanism based on the addition of the difference of two independent Poisson random variables as noise. According to their findings, even when the precision of the Skellam mechanism is low, it provides the same privacy-accuracy trade-off as the continuous Gaussian mechanism. In Bao *et al.* [57], they propose a Skellam mixture mechanism (SMM) based on injecting random noise from a mixture of two shifted symmetric Skellam distributions. The SMM is found to satisfy the $(\varepsilon - \delta)$ differential privacy condition. By

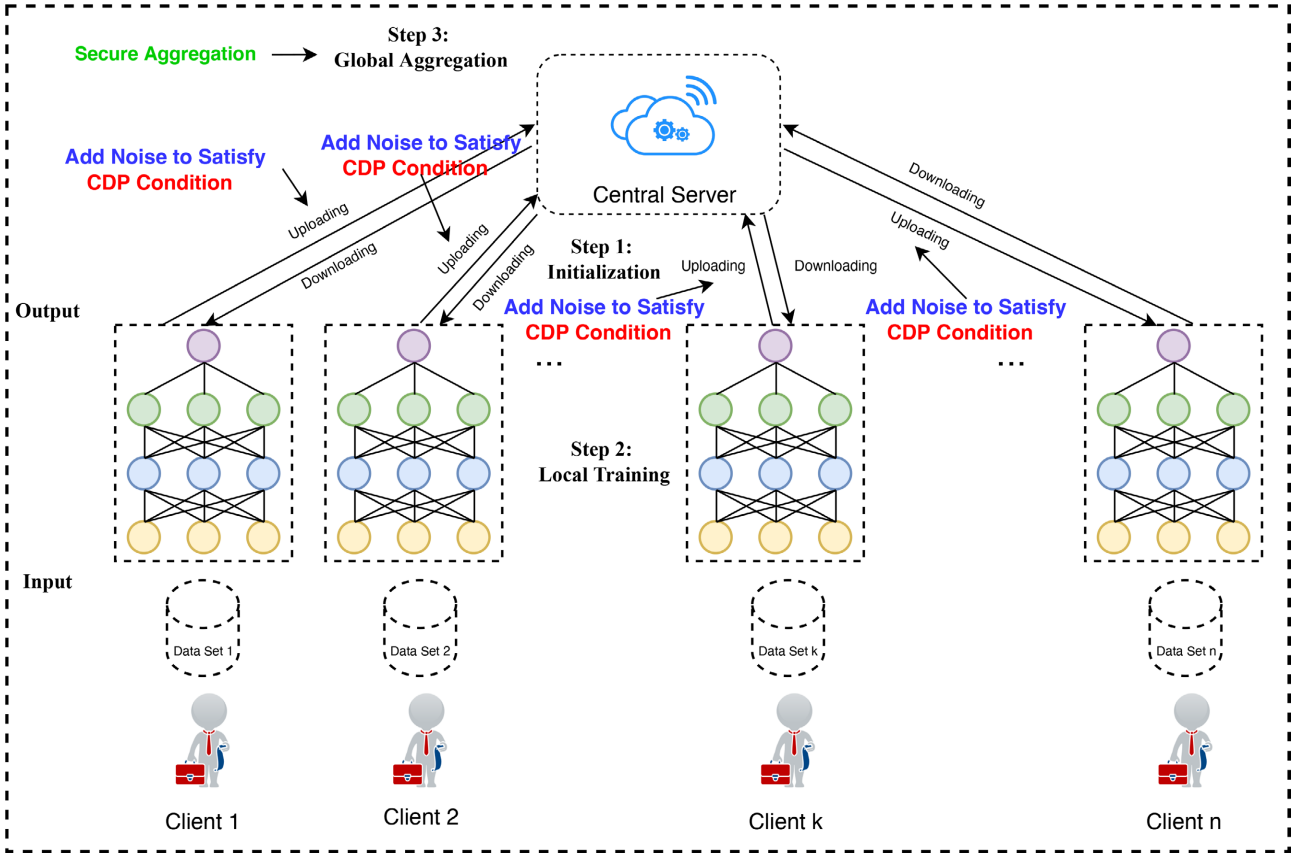

Figure 4. Workflow of federated learning with distributed differential privacy.

applying SMM to federated learning with distributed SGD, the authors show that SMM improves model utility by eliminating the step of rounding the gradients. Chen *et al.* [58] proposes a Poisson binomial mechanism (PBM) under which it encodes local information as parameters of a binomial distribution, resulting in discrete outputs. Theoretically results show that PBM satisfies the $(\varepsilon\text{-}\delta)$-approximate differential privacy, the communication cost equals is $O\left(d\left(\log_{10}\left(n+\left(d\varepsilon_{\mathrm{Dp}}^2\right)/\left(n\log_{10}\left(1/\delta\right)\right)\right)\right)\right)$ and an MSE is at most $O\left(\left(c^2d\right)/\left(n^2\varepsilon_{\mathrm{DP}}^2\right)\right)$. Chen *et al.* [59] characterize the fundamental communication cost required to obtain the best accuracy achievable under ε central differential privacy. Theoretical results show that $O\left(\min\left(n^2\varepsilon^2,d\right)\right)$ bits per client are both sufficient and necessary for obtaining the best accuracy achievable ε-differential privacy. Cheu *et al.* [60] investigate the Shuffling method in distributed differential privacy and show that this model provides the power of the central model while avoiding the need to trust a central server and the complexity of cryptographic secure function evaluation. Jiang *et al.* [61] focus on the client dropout problem in distributed differential privacy and propose a distributed differentially private FL framework, named Hyades. Results show that Hyades is capable of managing client dropout in various realistic scenarios and achieving the optimal privacy-utility trade-off. Table 3 summarizes the recent contributions in distributed differential privacy.

Table 3. Summary of contributions in distributed differential privacy.

Approach	Technique	Main Idea
Dwork *et al.* [52]	Binomial mechanism	Propose a Binomial mechanism.
Agarwal *et al.* [53]	FL + Binomial mechanism	Apply the binomial mechanism to federated learning and propose a stochastic *k*-level quantization method and a randomized rotation method to optimize the communication efficiency.
Canonne *et al.* [54]	Discrete Gaussian mechanism	Propose a discrete Gaussian mechanism.
Kairouz *et al.* [55]	FL + Discrete Gaussian mechanism	Apply the discrete Gaussian mechanism to federated learning, and results show that the model can provide $1/2\varepsilon^2$ central differential privacy.
Agarwal *et al.* [56]	FL + Skellam mechanism	Proposes a Skellam mechanism based on the addition of the difference of two independent Poisson random variables; Apply it to federated learning and achieve $(\alpha, \varepsilon(\alpha))$ Rényi differential privacy.
Bao *et al.* [57]	FL + Skellam mixture mechanism	Proposes a Skellam mixture mechanism based on a mixture of two shifted symmetric Skellam distributions; Apply it to federated learning and achieve $(\varepsilon - \delta)$ differential privacy condition.
Chen *et al.* [58]	FL + Poisson binomial mechanism	Proposes a Poisson binomial mechanism; Apply it to federated learning and achieve the $(\varepsilon - \delta)$-approximate differential privacy.

4. Optimization Techniques in Federated Learning with Differential Privacy

Applying differential privacy techniques to federation learning models can effectively protect clients' data security, however, differential privacy techniques also lead to problems such as decreasing model accuracy and increasing communication costs. Therefore, scholars have started to optimize the federated learning models with differential privacy, and the main directions of optimization are algorithm accuracy optimization and communication cost optimization.

Zhou and Tang [62] design a differentially private distributed algorithm based on the stochastic variance reduced gradient (SVRG) algorithm, which is capable of preventing the learning server from accessing and inferring private training data. The authors further quantify its impact on learning in terms of convergence rate and shows that noise added at each gradient update results in a bounded deviation from the optimal way of learning. Hu *et al.* [63] address the issue of privacy-preserving techniques for federated learning models under heterogeneous customer data sets. They show that the model satisfies $(\varepsilon - \delta)$ differential privacy when the Gaussian mechanism is used by each client. Van Dijk *et al.* [64] propose a new algorithm for asynchronous federated learning that eliminates

waiting times while at the same time reducing overall network communication. By adding Gaussian noise, they demonstrate how our algorithm can be made differentially private. Girgis *et al.* [65] focuses on the stochastic gradient descent algorithm for solving federated learning models with local differential privacy and proposes a distributed communication-efficient and locally differentially private stochastic gradient descent algorithm (CLDP-SGD) along with a detailed analysis of its communication, privacy, and convergence tradeoffs. Zhang *et al.* [32] examine the impact of clipping on federated learning with differential privacy and provide a convergence analysis of a differential private (DP) FedAvg algorithm. Zhang *et al.* [66] propose a federated learning scheme based on differential privacy and mechanism design. In addition to differential privacy mechanisms, two dominant-strategy truthful, individually rational, and budget-balanced mechanisms are designed to motivate clients to participate in training. Experiments demonstrate the effectiveness of the proposed scheme. Using optimal private linear operators on adaptive streams, Denisov *et al.* (2022) [67] present an improved Differential Privacy for SGD. The proposed algorithm achieves significant improvements in a notable problem in federated learning with differential privacy at the user level.

Lian *et al.* [68] presented COFEL, a novel federated learning system that reduces communication time through layer-based parameter selection and enhances privacy protection through local differences in privacy. In addition, they propose the COFEL-AVG algorithm for global aggregation as well as a layer-based parameter selection method, which enables the selection of the most valuable parameters for global aggregation in order to optimize the communication and training process. Amiri *et al.* [69] discuss the communication costs brought about by differential privacy, and present a novel algorithm for compressing client-server communications through quantization in order to achieve both differential privacy and reduced communication overhead. Liu *et al.* [70] propose a Projected Federated Averaging (PFA) scheme to explicitly model and leverage the heterogeneous privacy requirements of different clients and optimize utility for the joint model while minimizing communication cost. Truex *et al.* [71] present a novel approach that combines differential privacy and SMC, thus enabling users to reduce the growth of noise injection as the number of parties increases without sacrificing privacy while maintaining a pre-defined rate of trust. **Table 4** summarizes the optimization techniques in differential private federated learning models.

5. Application of Differential Private Federated Learning

Differential privacy-preserving federated learning techniques can effectively address data security issues in federated learning, and thus have achieved important applications in many fields.

Andrés *et al.* [72] investigate the application of differential privacy techniques in protecting customer geolocation data. They present a mechanism for achieving

Table 4. Summary of contributions in optimization techniques.

Approach	Technique	Main Idea
Zhou and Tang [62]	SVGR Algorithm	Develop a novel differentially private distributed algorithm based on the stochastic variance reduced gradient technique.
Van Dijk *et al.* [64]	Asynchronous federated learning algorithm	Develop a novel algorithm that eliminates waiting times and reduces overall network communication.
Girgis *et al.* [65]	CLDP-SGD	Develop a distributed communication-efficient and locally differentially private stochastic gradient descent algorithm along with a detailed analysis of its communication, privacy, and convergence tradeoffs.
Zhang *et al.* [66]	Mechanism Design	Develop a federated learning scheme based on differential privacy and mechanism design under which high-quality clients are selected to improve the accuracy of the model.
Denisov *et al.* [67]	Optimal Private Linear Operators	Develop improved differential privacy for SGD that achieves significant improvements in a notable problem in federated learning with differential privacy at the user level.
Lian *et al.* [68]	COFEL	Develop a novel federated learning system that reduces communication time through layer-based parameter selection and enhances privacy protection through local differences in privacy.
Amiri *et al.* [69]	Universal Vector Quantization	Develop a novel algorithm for achieving differential privacy and reduced communication overhead through compression of client-server communication by quantization.
Liu *et al.* [70]	FL-PFA	Develop a novel framework, named FL-PFA, that achieves communication cost minimization.
Zhang *et al.* [32]	Clipping	Develop a novel federated learning framework with clipping technique
Truex *et al.* [71]	Security Multi-party Computation	Develop a novel approach that combines differential privacy and SMC, thus enabling users to reduce the growth of noise injection.

geo-indistinguishability by adding controlled random noise to the user's location. Wang *et al.* [73] consider local differential privacy protection for both qualitative data (e.g., categorical data) and discrete quantitative data (e.g., location data). They derive a *k*-subset mechanism and an efficient extension of *k*-subset mechanism for categorical data and discrete quantitive data, respectively. Zhao *et al.* [74] study the application of federated learning in the Internet of Vehicles. In this context, user data, such as traffic information, vehicle registration information, etc., may be exposed. For this purpose, the authors propose a novel local

differential privacy mechanism, named as Three-Outputs, to protect the privacy of clients' data, and propose an LDP-FedSGD to train the model. Cao *et al.* [75] examine the application of differential private federated learning in the context of the Power Internet of Things. The authors propose IFed, a novel federated learning framework that takes into account the trade-off between local differential privacy, data utility, and resource consumption, to allow electric providers who normally have adequate computing resources to assist users in the Power Internet of Things. Jia *et al.* [76] propose a blockchain-enabled differential private federated learning in the Industrial Internet of Things (IIoT). Extensive experimental results show that the proposed scheme and working mechanism have better performance in the selected indicators. Olowononi *et al.* [77] propose the use of FL, together with differential privacy to improve the resiliency of vehicular cyber-physical systems to adversarial attacks in connected vehicles. Liu *et al.* [78] propose a federated learning framework for distributed medical institutions to collaboratively learn a prediction model. In comparison with state-of-the-art and in-depth ablation experiments, the proposed method performs better on two medical image segmentation tasks. Kaissis *et al.* [79] present PriMIA, a differential private federated learning framework for image analysis, and theoretically and empirically evaluate its performance and privacy guarantees, and demonstrate that the protections provided prevent gradient-based model inversion attacks from regenerating usable data. Adnan *et al.* (2022) [80] investigates the application of differentially private federated learning to the analysis of histopathology images. In a comparison of the performance of the conventional machine learning model with the federated learning model, the authors found that the federated learning model could achieve a similar performance while providing strong privacy guarantees. Zhang *et al.* (2022) [81] investigated the application of differential private federated learning models to industrial cyber-physical systems. The authors propose a Privacy-Enhanced Momentum Federated Learning framework called PEMFL, which incorporates differential privacy (DP), momentum federated learning (MFL) and chaos-based encryption methods. Theoretical analysis and experimental results demonstrate the excellent accuracy and privacy security of the PEMFL. Liu *et al.* [82] investigate the application of differential private federated learning to wireless sensor networks. As a result of integrating hybrid differential privacy into federated learning, the authors propose a secure and reliable federated learning algorithm. Based on a theoretical analysis and an experimental evaluation on real-world datasets, the validity of the algorithm is demonstrated. **Table 5** summarizes the applications of differential private federated learning models.

6. Future Directions

Differential privacy techniques have had some success in federated learning. Enterprises such as Microsoft, Apple, and Google have applied differential privacy-preserving federated learning models to their operations (Cormode *et al.* [83]).

Table 5. Summary of contributions in applications of DPFL.

Approach	Area	Main Idea
Andrés *et al.* [72]	Geography	Investigate the application of differential privacy techniques in protecting customer geolocation data.
Zhao *et al.* [74]	Internet of Vehicles	Propose a novel local differential privacy mechanism, named as Three-Outputs, to protect the privacy of client's data, and propose an LDP-FedSGD to train the model.
Cao *et al.* [75]	Power Internet of Things	Propose IFed, a novel federated learning framework that takes into account the trade-off between local differential privacy, data utility, and resource consumption, to allow electric providers who normally have adequate computing resources to assist users in the Power Internet of Things.
Jia *et al.* [76]	Industrial Internet of Things	Propose a blockchain-enabled differential private federated learning in Industrial Internet of Things (IIoT).
Olowononi *et al.* [77]	Vehicular Cyber-physical Systems	Propose a differential-private federated learning framework to improve the resiliency of vehicular cyber-physical systems to adversarial attacks in connected vehicles.
Liu *et al.* [78]	Medical Institutions	Propose a federated learning framework for distributed medical institutions to collaboratively learn a prediction model.
Kaissis *et al.* [79]	Medical Image Analysis	Propose a differential private federated learning framework for image analysis, named PriMIA, and theoretically and empirically evaluate its performance and privacy guarantees.
Liu *et al.* [82]	Wireless Sensor Networks	Propose a secure and reliable federated learning algorithm for wireless sensor networks.

But existing research is still lacking. In this section, we discuss three possible future research directions for differential privacy techniques in federation learning: research on the conditions for the use of differential privacy, research on the design of differential privacy features-based algorithms, and research on the combination of game theory and distributed differential privacy.

6.1. Research on the Conditions for the Use of Differential Privacy

Most of the existing studies on differential privacy-preserving federated learning assume that the central server or federated learning clients use differential privacy techniques from the beginning of model training, and there is no discussion on the conditions for using differential privacy techniques. We know that attacking a federated learning model is costly, so even a malicious individual will make a cost-benefit tradeoff before launching an attack on a federated learning model. If the cost of attacking a federated learning model is too high, *i.e.*, more

than the benefit from attacking a federated learning model, then a malicious individual will not launch an attack on the model. Thus, there is no need to use differential privacy techniques in a federated learning model under such conditions. Therefore, it is necessary to analyze the conditions under which an attacker initiates an attack from the perspective of the attacker's utility and thus determine the conditions under which differential privacy techniques should be used.

6.2. Research on the Design of Differential Privacy Features-Based Algorithms

The stochastic gradient descent algorithm is a common algorithm for solving large-scale differential privacy-preserving federated learning models. Scholars have made a series of optimizations on the convergency and convergence speed of the stochastic gradient descent algorithm. However, in some cases, the results of using the stochastic gradient descent algorithm to solve differential privacy-preserving federated learning models still fail to meet the requirements for industrial use. One possible reason is that the stochastic gradient descent algorithm is a general algorithm, and the characteristics of the differential privacy-preserving federated learning model are not fully considered in the design of this algorithm. Therefore, it is necessary to develop algorithms with better convergence and higher accuracy based on the features of differential privacy-preserving federated learning models.

6.3. Research on the Combination of Game Theory and Distributed Differential Privacy

Distributed differential privacy can improve the accuracy of federated learning models while protecting the security of federated learning clients' data. However, distributed differential privacy requires the use of secure aggregation techniques, which imposes expensive communication costs on the federated learning models. Therefore, federated learning models with distributed differential privacy usually include the step of federated learning client selection. Existing client selection methods are mainly based on probability theory in which the probability of a federated learning client being selected is constructed by the norm of the gradients uploaded by federated learning clients. This portrayal is too simple and does not sufficiently consider the contribution of federated learning clients to the federated learning model. A more reasonable way is to apply cooperative game theory and mechanism design theory to consider the game relationship between the central server and the federated clients and among the federated clients to make a reasonable portrayal of the contribution of the federated learning users, so as to select higher quality federated clients to participate in the federated learning model. Therefore, it is essential to conduct research that combines game theory with distributed differential privacy.

7. Conclusion

Differential privacy techniques are a key design element of federated learning

systems. In this work, we extensively survey state-of-the-art approaches and open up some interesting future research directions. First, we introduce the workflow of federated learning and the theoretical foundations of differential privacy techniques. Then, we overview three paradigms arising from the combination of differential privacy techniques and federation learning models, namely: centralized differential privacy, local differential privacy, and distributed differential privacy. After this, we review the optimization study of federated learning models oriented to differential privacy preservation. Finally, we review the applications of differential privacy-preserving federated learning models in various domains. In conclusion, differential privacy techniques play a crucial role in federated learning systems. From this survey, we expect more and more researchers to devote themselves to this field.

Acknowledgements

The authors would like to thank the editors and anonymous reviewers for their constructive comments, which help improve the study significantly. This work was supported by the National Natural Science Foundation of China [grant number 72201022].

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Bishop, C.M. and Nasrabadi, N.M. (2006) Pattern Recognition and Machine Learning. Springer, New York.

[2] Nadkarni, P.M., Ohno-Machado, L. and Chapman, W.W. (2011) Natural Language Processing: An Introduction. *Journal of the American Medical Informatics Association*, **18**, 544-551. https://doi.org/10.1136/amiajnl-2011-000464

[3] Jarvis, R.A. (1983) A Perspective on Range Finding Techniques for Computer Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-5**, 122-139. https://doi.org/10.1109/TPAMI.1983.4767365

[4] Fatima, M. and Pasha, M. (2017) Survey of Machine Learning Algorithms for Disease Diagnostic. *Journal of Intelligent Learning Systems and Applications*, **9**, 1-16. https://doi.org/10.4236/jilsa.2017.91001

[5] Bolton, R.J. and Hand, D.J. (2002) Statistical Fraud Detection: A Review. *Statistical Science*, **17**, 235-255. https://doi.org/10.1214/ss/1042727940

[6] Zhao, W., Chellappa, R., Phillips, P.J. and Rosenfeld, A. (2003) Face Recognition: A Literature Survey. *ACM Computing Surveys*, **35**, 399-458. https://doi.org/10.1145/954339.954342

[7] Reddy, D.R. (1976) Speech Recognition by Machine: A Review. *Proceedings of the IEEE*, **64**, 501-531. https://doi.org/10.1109/PROC.1976.10158

[8] Ji, Z., Lipton, Z.C. and Elkan, C. (2014) Differential Privacy and Machine Learning: A Survey and Review.

[9] El Ouadrhiri, A. and Abdelhadi, A. (2022) Differential Privacy for Deep and Fede-

rated Learning: A Survey. *IEEE Access*, **10**, 22359-22380. https://doi.org/10.1109/ACCESS.2022.3151670

[10] McMahan, B., Moore, E., Ramage, D., Hampson, S. and Arcas, B.A. (2017) Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the* 20*th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, 20-22 April 2017, 1273-1282.

[11] Ziller, A., Trask, A., Lopardo, A., *et al.* (2021) Pysyft: A Library for Easy Federated Learning. In: ur Rehman, M.H. and Gaber, M.M., Eds., *Federated Learning Systems*, Springer, Berlin, 111-139. https://doi.org/10.1007/978-3-030-70604-3_5

[12] Liu, Y., Fan, T., Chen, T., Xu, Q. and Yang, Q. (2021) FATE: An Industrial Grade Platform for Collaborative Learning with Data Protection. *Journal of Machine Learning Research*, **22**, 1-6.

[13] Li, Q., Wen, Z., Wu, Z., *et al.* (2021) A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2021.3124599

[14] Nasr, M., Shokri, R. and Houmansadr, A. (2019) Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-Box Inference Attacks against Centralized and Federated Learning. 2019 *IEEE Symposium on Security and Privacy*, San Francisco, 19-23 May 2019, 739-753. https://doi.org/10.1109/SP.2019.00065

[15] Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A. and Smith, V. (2020) Federated Optimization in Heterogeneous Networks. *Proceedings of Machine Learning and Systems*, Vol. 2, 429-450.

[16] Reisizadeh, A., Mokhtari, A., Hassani, H., Jadbabaie, A. and Pedarsani, R. (2020) Fedpaq: A Communication-Efficient Federated Learning Method with Periodic Averaging and Quantization. *International Conference on Artificial Intelligence and Statistics*, 26-28 August 2020, 2021-2031.

[17] So, J., Géler, B. and Avestimehr, A.S. (2021) Turbo-Aggregate: Breaking the Quadratic Aggregation Barrier in Secure Federated Learning. *IEEE Journal on Selected Areas in Information Theory*, **2**, 479-489. https://doi.org/10.1109/JSAIT.2021.3054610

[18] Wang, H., Yurochkin, M., Sun, Y., Papailiopoulos, D. and Khazaeni, Y. (2020) Federated Learning with Matched Averaging.

[19] Liu, L., Zhang, J., Song, S.H. and Letaief, K.B. (2020) Client-Edge-Cloud Hierarchical Federated Learning. *ICC* 2020 *IEEE International Conference on Communications*, Dublin, 7-11 June 2020, 1-6. https://doi.org/10.1109/ICC40277.2020.9148862

[20] Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T. and Yu, H. (2019) Federated Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, Vol. 13, Springer, Berlin, 1-207. https://doi.org/10.1007/978-3-031-01585-4

[21] Ur Rehman, M.H. and Gaber, M.M. (2021) Federated Learning Systems: Towards Next-Generation AI. Springer Nature, Berlin. https://doi.org/10.1007/978-3-030-70604-3

[22] Ludwig, H. and Baracaldo, N. (2022) Federated Learning: A Comprehensive Overview of Methods and Applications. Springer Nature, Berlin. https://doi.org/10.1007/978-3-030-96896-0

[23] Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006) Calibrating Noise to Sensitivity in Private Data Analysis. In: *Theory of Cryptography Conference*, Springer, Berlin, 265-284. https://doi.org/10.1007/11681878_14

[24] McSherry, F. and Talwar, K. (2007) Mechanism Design via Differential Privacy.

48*th Annual IEEE Symposium on Foundations of Computer Science*, Providence, 21-23 October 2007, 94-103. https://doi.org/10.1109/FOCS.2007.66

[25] Nikolov, A., Talwar, K. and Zhang, L. (2013) The Geometry of Differential Privacy: The Sparse and Approximate Cases. *Proceedings of the* 45*th Annual ACM Symposium on Theory of Computing*, Palo Alto, 2-4 June 2013, 351-360. https://doi.org/10.1145/2488608.2488652

[26] Dwork, C. (2008) Differential Privacy: A Survey of Results. *International Conference on Theory and Applications of Models of Computation*, Xi'an, 25-29 April 2008, 1-19. https://doi.org/10.1007/978-3-540-79228-4_1

[27] Dwork, C. and Roth, A. (2014) The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science*, **9**, 211-407. https://doi.org/10.1561/0400000042

[28] Geyer, R.C., Klein, T. and Nabi, M. (2017) Differentially Private Federated Learning: A Client Level Perspective.

[29] Triastcyn, A. and Faltings, B. (2019) Federated Learning with Bayesian Differential Privacy. 2019 *IEEE International Conference on Big Data*, Los Angeles, 9-12 December 2019, 2587-2596. https://doi.org/10.1109/BigData47090.2019.9005465

[30] Wei, K., Li, J., Ding, M., Ma, C., *et al.* (2020) Federated Learning with Differential Privacy: Algorithms and Performance Analysis. *IEEE Transactions on Information Forensics and Security*, **15**, 3454-3469. https://doi.org/10.1109/TIFS.2020.2988575

[31] Bernau, D., Robl, J., Grassal, P.W., Schneider, S. and Kerschbaum, F. (2021) Comparing Local and Central Differential Privacy Using Membership Inference Attacks. *IFIP Annual Conference on Data and Applications Security and Privacy*, Calgary, 19-20 July 2021, 22-42. https://doi.org/10.1007/978-3-030-81242-3_2

[32] Zhang, X., Chen, X., Hong, M., Wu, S. and Yi, J. (2022) Understanding Clipping for Federated Learning: Convergence and Client-Level Differential Privacy. *International Conference on Machine Learning*, Baltimore, 17-23 July 2022, 26048-26067.

[33] Hu, R., Gong, Y. and Guo, Y. (2022) Federated Learning with Sparsified Model Perturbation: Improving Accuracy under Client-Level Differential Privacy.

[34] Li, Z., Huang, Z., Chen, C. and Hong, C. (2019) Quantification of the Leakage in Federated Learning.

[35] Melis, L., Song, C., De Cristofaro, E. and Shmatikov, V. (2019) Exploiting Unintended Feature Leakage in Collaborative Learning. 2019 *IEEE Symposium on Security and Privacy*, San Francisco, 19-23 May 2019, 691-706. https://doi.org/10.1109/SP.2019.00029

[36] Kasiviswanathan, S.P., Lee, H.K., Nissim, K., Raskhodnikova, S. and Smith, A. (2011) What Can We Learn Privately? *SIAM Journal on Computing*, **40**, 793-826. https://doi.org/10.1137/090756090

[37] Erlingsson, Ú., Pihur, V., Korolova, A. (2014) Rappor: Randomized Aggregatable Privacy-Preserving Ordinal Response. *Proceedings of the* 2014 *ACM SIGSAC Conference on Computer and Communications Security*, Scottsdale, 3-7 November 2014, 1054-1067. https://doi.org/10.1145/2660267.2660348

[38] Liu, F. (2018) Generalized Gaussian Mechanism for Differential Privacy. *IEEE Transactions on Knowledge and Data Engineering*, **31**, 747-756. https://doi.org/10.1109/TKDE.2018.2845388

[39] Truex, S., Liu, L., Chow, K.H., Gursoy, M.E. and Wei, W. (2020) LDP-Fed: Federated Learning with Local Differential Privacy. *Proceedings of the* 3*rd ACM International Workshop on Edge Systems, Analytics and Networking*, Heraklion, 27 April 2020, 61-66. https://doi.org/10.1145/3378679.3394533

[40] Sun, Z., Kairouz, P., Suresh, A.T. and McMahan, H.B. (2019) Can You Really Backdoor Federated Learning?

[41] Xu, M., Ding, B., Wang, T. and Zhou, J. (2020) Collecting and Analyzing Data Jointly from Multiple Services under Local Differential Privacy. *Proceedings of the VLDB Endowment*, **13**, 2760-2772. https://doi.org/10.14778/3407790.3407859

[42] Naseri, M., Hayes, J. and De Cristofaro, E. (2020) Toward Robustness and Privacy in Federated Learning: Experimenting with Local and Central Differential Privacy.

[43] Wang, Y., Tong, Y. and Shi, D. (2020) Federated Latent Dirichlet Allocation: A Local Differential Privacy based Framework. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 6283-6290. https://doi.org/10.1609/aaai.v34i04.6096

[44] Girgis, A., Data, D. and Diggavi, S. (2021) Renyi Differential Privacy of the Subsampled Shuffle Model in Distributed Learning. *Advances in Neural Information Processing Systems*, Vol. 34, 29181-29192. https://doi.org/10.1145/3460120.3484794

[45] Wei, K., Li, J., Ding, M., Ma, C., Su, H., Zhang, B. and Poor, H.V. (2022) User-Level Privacy-Preserving Federated Learning: Analysis and Performance Optimization. *IEEE Transactions on Mobile Computing*, **21**, 3388-3401. https://doi.org/10.1109/TMC.2021.3056991

[46] Zhou, H., Yang, G., Dai, H. and Liu, G. (2022) PFLF: Privacy-Preserving Federated Learning Framework for Edge Computing. *IEEE Transactions on Information Forensics and Security*, **17**, 1905-1918. https://doi.org/10.1109/TIFS.2022.3174394

[47] Thapa, C., Arachchige, P.C.M., Camtepe, S. and Sun, L. (2022) Splitfed: When Federated Learning Meets Split Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 8485-8493. https://doi.org/10.1609/aaai.v36i8.20825

[48] Wu, C., Wu, F., Lyu, L., Qi, T., Huang, Y. and Xie, X. (2022) A Federated Graph Neural Network Framework for Privacy-Preserving Personalization. *Nature Communications*, **13**, Article No. 3091. https://doi.org/10.1038/s41467-022-30714-9

[49] Wang, C., Wu, X., Liu, G., Deng, T., Peng, K. and Wan, S. (2022) Safeguarding Cross-Silo Federated Learning with Local Differential Privacy. *Digital Communications and Networks*, **8**, 446-454. https://doi.org/10.1016/j.dcan.2021.11.006

[50] Zheng, Q., Chen, S., Long, Q. and Su, W. (2021) Federated F-Differential Privacy. *International Conference on Artificial Intelligence and Statistics*, 13-15 April 2021, 2251-2259.

[51] Bonawitz, K., Ivanov, V., Kreuter, B., *et al.* (2017) Practical Secure Aggregation for Privacy-Preserving Machine Learning. *Proceedings of the* 2017 *ACM SIGSAC Conference on Computer and Communications Security*, Dallas, 30 October-3 November 2017, 1175-1191. https://doi.org/10.1145/3133956.3133982

[52] Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I. and Naor, M. (2006) Our Data, Ourselves: Privacy via Distributed Noise Generation. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, St. Petersburg, 28 May-1 June 2006, 486-503. https://doi.org/10.1007/11761679_29

[53] Agarwal, N., Suresh, A.T., Yu, F.X.X., Kumar, S. and McMahan, B. (2018) cpSGD: Communication-Efficient and Differentially-Private Distributed SGD. *Advances in Neural Information Processing Systems*, Vol. 31, 1-12.

[54] Canonne, C.L., Kamath, G. and Steinke, T. (2020) The Discrete Gaussian for Differential Privacy. *Advances in Neural Information Processing Systems*, Vol. 33, 15676-15688.

[55] Kairouz, P., Liu, Z. and Steinke, T. (2021) The Distributed Discrete Gaussian Mechanism for Federated Learning with Secure Aggregation. *International Conference on Machine Learning*, 18-24 July 2021, 5201-5212.

[56] Agarwal, N., Kairouz, P. and Liu, Z. (2021) The Skellam Mechanism for Differentially Private Federated Learning. *Advances in Neural Information Processing Systems*, Vol. 34, 5052-5064.

[57] Bao, E., Zhu, Y., Xiao, X., Yang, Y., Ooi, B.C., Tan, B.H.M. and Aung, K.M.M. (2022) Skellam Mixture Mechanism: A Novel Approach to Federated Learning with Differential Privacy. *Proceedings of the VLDB Endowment*, **15**, 2348-2360. https://doi.org/10.14778/3551793.3551798

[58] Chen, W.N., Ozgur, A. and Kairouz, P. (2022) The Poisson Binomial Mechanism for Unbiased Federated Learning with Secure Aggregation. *International Conference on Machine Learning*, Baltimore, 17-23 July 2022, 3490-3506.

[59] Chen, W.N., Choo, C.A.C., Kairouz, P. and Suresh, A.T. (2022) The Fundamental Price of Secure Aggregation in Differentially Private Federated Learning. *International Conference on Machine Learning*, Baltimore, 17-23 July 2022, 3056-3089.

[60] Cheu, A., Smith, A., Ullman, J., Zeber, D. and Zhilyaev, M. (2019) Distributed Differential Privacy via Shuffling. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Darmstadt, 19-23 May 2019, 375-403. https://doi.org/10.1007/978-3-030-17653-2_13

[61] Jiang, Z., Wang, W. and Chen, R. (2022) Taming Client Dropout for Distributed Differential Privacy in Federated Learning.

[62] Zhou, Y. and Tang, S. (2020) Differentially Private Distributed Learning. *INFORMS Journal on Computing*, **32**, 779-789. https://doi.org/10.1287/ijoc.2019.0912

[63] Hu, R., Guo, Y., Li, H., Pei, Q. and Gong, Y. (2020) Personalized Federated Learning with Differential Privacy. *IEEE Internet of Things Journal*, **7**, 9530-9539. https://doi.org/10.1109/JIOT.2020.2991416

[64] Van Dijk, M., Nguyen, N.V., Nguyen, T.N., Nguyen, L.M., Tran-Dinh, Q. and Nguyen, P.H. (2020) Asynchronous Federated Learning with Reduced Number of Rounds and with Differential Privacy from Less Aggregated Gaussian Noise.

[65] Girgis, A., Data, D., Diggavi, S., Kairouz, P. and Suresh, A.T. (2021) Shuffled Model of Differential Privacy in Federated Learning. *International Conference on Artificial Intelligence and Statistics*, 13-15 April 2021, 2521-2529.

[66] Zhang, L., Zhu, T., Xiong, P., Zhou, W. and Yu, P. (2022) A Robust Game-Theoretical Federated Learning Framework with Joint Differential Privacy. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2021.3140131

[67] Denisov, S., McMahan, B., Rush, K., Smith, A. and Thakurta, A. (2022) Improved Differential Privacy for SGD via Optimal Private Linear Operators on Adaptive Streams. *Advances in Neural Information Processing Systems*.

[68] Lian, Z., Wang, W. and Su, C. (2021) COFEL: Communication-efficient and Optimized Federated Learning with Local Differential Privacy. *ICC* 2021-*IEEE International Conference on Communications*, Montreal, 14-23 June 2021, 1-6. https://doi.org/10.1109/ICC42927.2021.9500632

[69] Amiri, S., Belloum, A., Klous, S. and Gommans, L. (2021) Compressive Differentially Private Federated Learning through Universal Vector Quantization. *AAAI Workshop on Privacy-Preserving Artificial Intelligence*, 2-9 February 2021, 1-5.

[70] Liu, J., Lou, J., Xiong, L., Liu, J. and Meng, X. (2021) Projected Federated Averaging with Heterogeneous Differential Privacy. *Proceedings of the VLDB Endowment*, **15**, 828-840. https://doi.org/10.14778/3503585.3503592

[71] Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. and Zhou, Y. (2019) A Hybrid Approach to Privacy-Preserving Federated Learning. *Proceedings*

of the 12*th ACM Workshop on Artificial Intelligence and Security*, London, 15 November 2019, 1-11. https://doi.org/10.1145/3338501.3357370

[72] Andr's, M.E., Bordenabe, N.E., Chatzikokolakis, K. and Palamidessi, C. (2013) Geo-Indistinguishability: Differential Privacy for Location-Based Systems. *Proceedings of the* 2013 *ACM SIGSAC Conference on Computer & Communications Security*, Berlin, 4-8 November 2013, 901-914. https://doi.org/10.1145/2508859.2516735

[73] Wang, S., Huang, L., Nie, Y., Zhang, X., Wang, P., Xu, H. and Yang, W. (2019) Local Differential Private Data Aggregation for Discrete Distribution Estimation. *IEEE Transactions on Parallel and Distributed Systems*, **30**, 2046-2059. https://doi.org/10.1109/TPDS.2019.2899097

[74] Zhao, Y., Zhao, J., Yang, M., *et al.* (2020) Local Differential Privacy-Based Federated Learning for Internet of Things. *IEEE Internet of Things Journal*, **8**, 8836-8853. https://doi.org/10.1109/JIOT.2020.3037194

[75] Cao, H., Liu, S., Zhao, R. and Xiong, X. (2020) IFed: A Novel Federated Learning Framework for Local Differential Privacy in Power Internet of Things. *International Journal of Distributed Sensor Networks*, **16**. https://doi.org/10.1177/1550147720919698

[76] Jia, B., Zhang, X., Liu, J., Zhang, Y., Huang, K. and Liang, Y. (2021) Blockchain-Enabled Federated Learning Data Protection Aggregation Scheme with Differential Privacy and Homomorphic Encryption in IIoT. *IEEE Transactions on Industrial Informatics*, **18**, 4049-4058. https://doi.org/10.1109/TII.2021.3085960

[77] Olowononi, F.O., Rawat, D.B. and Liu, C. (2021) Federated Learning with Differential Privacy for Resilient Vehicular Cyber Physical Systems. 2021 *IEEE* 18*th Annual Consumer Communications & Networking Conference* (*CCNC*), Las Vegas, 9-12 January 2021, 1-5. https://doi.org/10.1109/CCNC49032.2021.9369480

[78] Liu, Q., Chen, C., Qin, J., Dou, Q. and Heng, P.A. (2021) Feddg: Federated Domain Generalization on Medical Image Segmentation via Episodic Learning in Continuous Frequency Space. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 1013-1023. https://doi.org/10.1109/CVPR46437.2021.00107

[79] Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., *et al.* (2021) End-to-End Privacy Preserving Deep Learning on Multi-Institutional Medical Imaging. *Nature Machine Intelligence*, **3**, 473-484. https://doi.org/10.1038/s42256-021-00337-8

[80] Adnan, M., Kalra, S., Cresswell, J.C., Taylor, G.W. and Tizhoosh, H.R. (2022) Federated Learning and Differential Privacy for Medical Image Analysis. *Scientific Reports*, **12**, Article No. 1953. https://doi.org/10.1038/s41598-022-05539-7

[81] Zhang, Z., Zhang, L., Li, Q., Wang, K., He, N. and Gao, T. (2022) Privacy-Enhanced Momentum Federated Learning via Differential Privacy and Chaotic System in Industrial Cyber-Physical Systems. *ISA Transactions*, **128**, 17-31. https://doi.org/10.1016/j.isatra.2021.09.007

[82] Liu, W., Cheng, J., Wang, X., Lu, X. and Yin, J. (2022) Hybrid Differential Privacy based Federated Learning for Internet of Things. *Journal of Systems Architecture*, **124**, Article ID: 102418. https://doi.org/10.1016/j.sysarc.2022.102418

[83] Cormode, G., Jha, S., Kulkarni, T., Li, N., Srivastava, D. and Wang, T. (2018) Privacy at Scale: Local Differential Privacy in Practice. *Proceedings of the* 2018 *International Conference on Management of Data*, Houston, 10-15 June 2018, 1655-1658. https://doi.org/10.1145/3183713.3197390