Personalized Differentially Private Federated Learning without Exposing Privacy Budgets

Junxu Liu
Renmin University of China
junxu_liu@ruc.edu.cn

Jian Lou
Zhejiang University
jian.lou@zju.edu.cn

Li Xiong
Emory University
lxiong@emory.edu

Xiaofeng Meng
Renmin University of China
xfmeng@ruc.edu.cn

ABSTRACT

The meteoric rise of cross-silo Federated Learning (FL) is due to its ability to mitigate data breaches during collaborative training. To further provide rigorous privacy protection with consideration of the varying privacy requirements across different clients, a privacy-enhanced line of work on personalized differentially private federated learning (PDP-FL) has been proposed. However, the existing solution for PDP-FL [21] assumes the raw privacy budgets of all clients should be collected by the server. These values are then *directly* utilized to improve the model utility via facilitating the privacy preferences partitioning (i.e., partitioning all clients into multiple privacy groups). It is however non-realistic because the raw privacy budgets can be quite informative and sensitive.

In this work, our goal is to achieve PDP-FL without exposing clients' raw privacy budgets by indirectly partitioning the privacy preferences solely based on clients' noisy model updates. The crux lies in the fact that the noisy updates could be influenced by two entangled factors of DP noises and non-IID clients' data, leaving it unknown whether it is possible to uncover privacy preferences by disentangling the two affecting factors. To overcome the hurdle, we systematically investigate the unexplored question of *under what conditions can the model updates of clients be primarily influenced by noise levels rather than data distribution*. Then, we propose a simple yet effective strategy based on clustering the L^2 norm of the noisy updates, which can be integrated into the vanilla PDP-FL to maintain the same performance. Experimental results demonstrate the effectiveness and feasibility of our privacy-budget-agnostic PDP-FL method.

CCS CONCEPTS

• **Security and privacy → Privacy protections**.

KEYWORDS

Differential Privacy, Federated Learning, Personalized Privacy

CIKM '23, October 21–25, 2023, Birmingham, United Kingdom.

ACM ISBN 979-8-4007-0124-5/23/10…$15.00
https://doi.org/10.1145/3583780.3615247

ACM Reference Format:
Junxu Liu, Jian Lou, Li Xiong, and Xiaofeng Meng. 2023. Personalized Differentially Private Federated Learning without Exposing Privacy Budgets. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23), October 21–25, 2023, Birmingham, United Kingdom.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3583780.3615247

1 INTRODUCTION

Cross-silo Federated Learning (FL) [12, 25], which allows multiple clients to collaboratively train a global model without requiring access to clients' raw data, has been widely adopted both in academia and industry. Differential Privacy [5, 6] has been further integrated into FL, which gives rise to the DP-FL studies [2, 3, 7, 20, 22–24, 28, 29] that seek to provide mathematically rigorous privacy protection at the desired level quantified by the privacy budget (denoted by ε). DP-FL bears much resemblance to non-DP FL in training (e.g., by building on top of FedAvg [25]) but additionally incorporates local updates clipping and Gaussian noise injection [1, 4, 26, 32], whereby clients' local updates will be more strictly protected.

A more challenging yet practical problem is personalized differentially private federated learning[1](PDP-FL, see Definition 1), which takes the wide-ranging differences in individuals' privacy preferences [11, 27, 30] into consideration and enables clients to predefine their own privacy budgets (as opposed to shared an identical value specified by the server) [21]. One common way to achieve PDP in FL is to add different amounts of Gaussian noise to clients' submitted local updates, while directly aggregating the noisy and discordant local updates would inevitably lead to suboptimal model performance due to the biased estimation of the global parameters. To address these issues, Liu et al. [21] present the first promising attempt by developing a projection-based approach named projected federated averaging (PFA) for noise reduction [8, 34]. However, a major downside of PFA is that they *treat clients' privacy budgets as publicly available knowledge* and allow the server to utilize this information directly to identify the conservative/liberal clients at the initialization stage (see Line 5, Algorithm 1 in Section 2).

Definition 1 (Personalized Differential Privacy in Federated Learning [21]). Let the set of clients be $C = \{C_1, \ldots, C_M\}$, where each client $C_m \in C$ holds a local dataset $\mathcal{D}_m$. The federated learning

[1]Within the cross-silo FL we are considering, each *client*'s local dataset consists of multiple records gathered from different *users*. Each user contributes only one individual record to their respective client's dataset. The term "personalized" is used to characterize customized DP guarantees for each client.

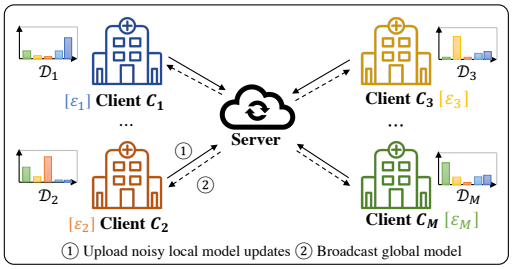


Figure 1: An illustration of the PDP-FL framework in which heterogeneous clients with non-IID data and personalized privacy budgets are collaboratively training a global model.

satisfies $\{(\varepsilon_m, \delta_m)\}_{m\in[M]}$-PDP, if each client C_m satisfies $(\varepsilon_m, \delta_m)$-DP with respect to its local dataset.

We argue that obtaining access to clients' privacy budgets can be problematic, especially in scenarios involving a server that is honest but curious. The raw value of clients' privacy budget is highly informative and sensitive, making it potentially a trigger for privacy attacks. In light of this, the primary objective is to discern the implicit privacy preferences of clients while keeping their privacy budgets undisclosed. More precisely, we focus on addressing this issue solely by utilizing the exchanged noisy local model updates between the server and the clients. In PFA, differences in privacy budgets among clients indicate that the amount of noise added to the gradients during training varies. Besides, the heterogeneous (non-IID) data also introduces a drift in the local updates of each client [13, 16, 19]. Thus, the crucial first step of indirectly estimating privacy preferences is to answer the following fundamental question: *under what conditions can the model updates of clients be primarily influenced by noise levels rather than data distribution*? To the best of our knowledge, there is currently no existing research that delves into this problem.

In summary, the contributions of this paper are twofold.

- We conduct a systematic evaluation using FedAvg and observe that the L^2-norms of local updates can effectively serve as an indicator for distinguishing between clients with diverse privacy budgets and non-IID data. This novel insight propels the advancement of the study on indirect privacy preferences partitioning.
- We introduce a simple yet effective approach for clustering clients, leveraging the off-the-shelf methods (e.g., Gaussian Mixture Models algorithm) without requiring any auxiliary knowledge about the real privacy budgets. Furthermore, we integrate the approach into PFA and validate its effectiveness through a series of comprehensive replicated experiments.

2 PRELIMINARIES

Federated Averaging (FedAvg). FedAvg [25] is the most widely used algorithm for solving the federated optimization problem. In each communication round, a randomly sampled subset of clients runs a certain number of Stochastic Gradient Descent (SGD) steps locally and independently, then the server averages the local updates and broadcasts a single global model to all clients. One of the limitations of FedAvg is its lack of special adjustments when encountering non-IID client data, resulting in suboptimal performance under such conditions [9, 17, 18].

Algorithm 1: Projected Federated Averaging with Indirect Privacy Preferences Partition

input : Number of clients M, clients' privacy budgets $(\varepsilon_1, \ldots, \varepsilon_M)$, DP parameter δ, number of communication rounds T, number of local steps τ
output: global model $\mathbf{x}_T$

Framework PDP-FL$((\varepsilon_1, \ldots, \varepsilon_M), \delta, T, \tau)$:
  // Partition clients into "public" and "private"
  $S_{pub}, S_{pri} \leftarrow$
  (Before) Direct partitioning based on exposed privacy budgets $(\varepsilon_1, \ldots, \varepsilon_M)$
  (After) Indirect partitioning based on clustering with L^2-norms of the noisy local updates $(\Delta\mathbf{x}_1, \ldots, \Delta\mathbf{x}_M)$ (the warm-start round)
  for *round* $t = 2, \ldots, T$ **do**
    $S^t \leftarrow$ (sample a random subset of $K < M$ clients)
    $S^t_{pub}, S^t_{pri} \leftarrow$ (partition the subset into "public" and "private")
    foreach $m \in S^t$ **do** *in parallel*
      $\Delta\mathbf{x}^t_m \leftarrow$ DPSGD$(t, \mathbf{x}^t, \tau)$
    (Before) $\Delta\bar{\mathbf{x}}^t \leftarrow$ PFA$(\{(\varepsilon_m, \Delta\mathbf{x}^t_m)\}_{m\in S^t}, S^t_{pub}, S^t_{pri})$
    (After) $\Delta\bar{\mathbf{x}}^t \leftarrow$ PFA$(\{\Delta\mathbf{x}^t_m\}_{m\in S^t}, S^t_{pub}, S^t_{pri})$
    $\mathbf{x}^{t+1} \leftarrow \mathbf{x}^t - \Delta\bar{\mathbf{x}}^t$
  return $\mathbf{x}^T$

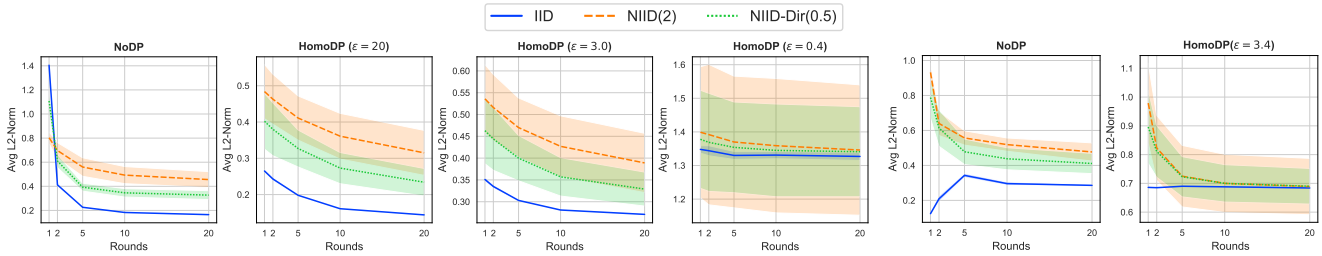
Function PFA$(\{\Delta\mathbf{x}_m\}_{m\in S}, S_{pub}, S_{pri})$:
  // Compute the subspace from "public" updates
  $\mathbf{V}_k \leftarrow$ (The top-k eigenvectors of the second moment matrix computed from all $\Delta\mathbf{x}_m$ and $m \in S_{pub}$)
  // Project "private" updates onto the subspace
  $\Delta\hat{\mathbf{x}}_{pri} \leftarrow \mathbf{V}_k\mathbf{V}_k^\top \frac{1}{|S_{pri}|}\sum_{m\in S_{pri}} \Delta\mathbf{x}^m$
  // Projected federated averaging
  $S \leftarrow S_{pub} + S_{pri}$
  $\Delta\bar{\mathbf{x}} \leftarrow \frac{|S_{pub}|}{|S|}\cdot\Delta\bar{\mathbf{x}}_{pub} + \frac{|S_{pri}|}{|S|}\cdot\Delta\hat{\mathbf{x}}_{pri}$
  return $\Delta\bar{\mathbf{x}}$

Differentially Private Federated Learning. While FL is effective at mitigating systemic privacy risks, it could disclose sensitive information through exchanged model updates derived from local data. To counter the potential privacy inference attacks from both the honest-but-curious server and malicious third parties, the integration of user-level Differential Privacy (DP) has become a common practice in the development of FL algorithms. A widely adopted technique involves introducing controlled Gaussian noise to clipped gradients during each local SGD iteration (i.e., DP-SGD [1]), resulting in a private version of FedAvg known as DP-FedAvg.

Projected Federated Averaging (PFA). In PFA [21], all clients are divided into two types according to their privacy budgets (i.e., "private" clients with stricter privacy budgets and "public" clients with more relaxed privacy budgets) exposed to the server at the initialization stage; then the server extracts a reduced-dimensional subspace from the "public" model updates and projects the "private" model updates onto it. In this way, the heavy private perturbation of the "private" updates can be discarded, and the most useful information from all clients can be aggregated to improve the joint model utility. Pseudocode is given in Algorithm 1.

3 STUDY ON THE IMPLICATIONS OF NOISE LEVEL AND DATA DISTRIBUTION

To answer the fundamental question posed in Section 1, in this section, we will empirically analyze the characteristics of the local model updates using the representative FedAvg algorithm under



(a) Experimental results evaluated on MNIST-LogR

(b) Experimental results evaluated on CIFAR10-CNN

Figure 2: The changes in the average and standard deviation of the L^2-norms of local updates (y-axis) across different communication rounds (x-axis) under various combinations of noise levels and data distributions.

various combinations of noise levels and data distributions. We consider two classic image classification tasks in the cross-silo FL setting: the MNIST [15] digit recognition with a simple logistic regression model (MNIST-LogR) and the CIFAR10 [14] image classification with the same CNN architecture as in [25] (CIFAR10-CNN).

3.1 Experimental Setup

Data distributions. To examine the effects of data heterogeneity, we first establish the baseline using IID data and consider two partitioning strategies to simulate potential non-IID scenarios.

- **NIID(2)**: a.k.a. the *quantity-based label distribution skew* where each client possesses data records associated with a fixed number (e.g., two) of distinct labels [25].
- **NIID-Dir(0.5)**: a.k.a. the *distribution-based label imbalance* where a $p_{k,m} \sim Dir(\beta)$ proportion of records of class k are allocated to client m. Here $Dir(\beta)$ denotes a Dirichlet distribution [10] and the smaller the β is, the resulting partition is more unbalanced. We choose the same $\beta = 0.5$ as done in [33].

Varying privacy budgets. We explore a diverse range of cumulative privacy budgets ε to observe the differences in local updates among clients with different privacy preferences (e.g., $\varepsilon \approx 0.4, 3.0, 20$[2] for the MNIST-LogR experiments).

Evaluation metrics. In this section, we always report the average and standard deviation of the L^2-norms of local updates across all clients (abbreviated to *avg./std. L^2-norms* for the sake of readability).

Hyperparameters. Unless otherwise stated, we fix the number of clients $M = 10$, the total number of rounds $T = 20$, the local minibatch size $B = 64$, the local steps $\tau = \lfloor |\mathcal{D}_m|/B \rfloor$ and the step size $\eta = 0.01$. The full participation paradigm ensures all clients get continuous observations throughout the training process.

3.2 Evaluation Results

For each plot in Fig. 2, we depict the changes in avg./std. L^2-norms across different communication rounds within the IID, NIID(2) and NIID-dir(0.5) settings, respectively. Additionally, we conduct a series of comparative experiments regarding varying levels of additive Gaussian noise. Note that in all experiments involving DP, we assume all clients share identical privacy budgets. This condition is referred to as homogeneous DP (HomoDP).

The isolated influence of data distribution. Two common trends can be observed from all plots of Fig. 2: (1) *Difference between IID and NIID*: both the avg. and the std. L^2-norms in IID cases consistently exhibit lower values compared to all non-IID cases along the training process; (2) *NIID-Dir(0.5) vs. NIID(2)*: in the majority of cases, NIID-Dir(0.5) tends to produce avg./std. L^2-norms that are either smaller or comparable to those obtained with NIID(2).

The isolated influence of noise level. From Fig. 2 (a), we can observe a clear negative correlation between the value of ε and the avg./std. L^2-norms in the two non-IID cases. Despite not being readily apparent, the observation remains consistent in the IID case. It makes sense since the discrepancies in privacy budgets imply different amounts of additive noise being introduced to the model updates during the local training procedure. Of particular interest is the resemblance in results between the cases with $\varepsilon = 3.0$ and $\varepsilon = 20$, indicating that both cases result in a comparable degree of perturbation on the magnitude of clients' local updates, despite the substantial gap in privacy budgets. In other words, when the value of ε gets smaller, the resulting enhancement in privacy becomes more noticeable. Given that similar trends have been observed in the CIFAR10-CNN experiments, we present only a partial set of results here due to the strict space limitations.

4 PDP-FL WITHOUT EXPOSING RAW PRIVACY BUDGETS

Now our focus shifts back to the PDP-FL setting where the additive Gaussian noises of the clients are drawn from different distributions determined by their privacy budgets. We propose a privacy-budget-agnostic adaptation of PFA and evaluate its effectiveness through experiments conducted under the identical setup as outlined in the prior PFA study.

4.1 Indirect Privacy Preferences Partitioning

Key Insight. In the previous section, we delved into how non-IID data and varying privacy budgets affect the local model updates of clients. Our experimental findings suggest the feasibility of indirectly partitioning the privacy preferences of clients into distinct

[2]It's worth mentioning that within current research on DP-ML, it's a common practice to conduct experiments with $\varepsilon > 1$ for better privacy-utility trade-offs. For example, [1] considers ε values of 2, 4 and 8, and [2] examines ε ranging from 1 to 20.

Table 1: Distribution of clients' privacy budgets

Distribution	Parameters Setting
MixGauss1	Mixture of $\mathcal{N}_1(0.1, 0.01)$ and $\mathcal{N}_2(10.0, 0.1)$ with mixture weights 0.9 and 0.1
MixGauss2	Mixture of $\mathcal{N}_1(1.0, 0.1)$ and $\mathcal{N}_2(10.0, 0.1)$ with mixture weights 0.9 and 0.1
MixGauss3	Mixture of $\mathcal{N}_1(0.1, 0.01)$, $\mathcal{N}_2(1.0, 0.1)$ and $\mathcal{N}_3(10.0, 1.0)$ with mixture weights 0.5, 0.4 and 0.1

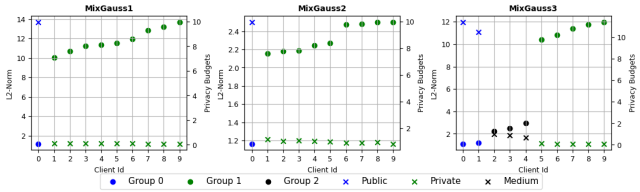


Figure 3: The consistency between the results of GMMs clustering based on L^2-norm (left y-axis) and the ground truths based on the real privacy budgets (right y-axis) across 10 clients (x-axis) evaluated on MNIST-LogR in NIID(2) setting.

groups by analyzing solely the L^2-norms of their local noisy updates, as long as the "private" clients choose privacy budgets that are sufficiently small (e.g., $\varepsilon < 1$) to ensure effective differentiation from the "public" ones.

Proposed Approach. In practical scenarios, it is reasonable to assume that conservative clients would prefer choosing stricter privacy budgets, while liberal ones might adopt more relaxed values. Besides, clients who have comparable privacy budgets are expected to introduce Gaussian noises drawn from a similar random distribution. This distinction in values of ε among different groups aligns with the condition emphasized in the aforementioned key insight, which inspires us to devise a clustering-based approach based on the L^2-norms. More specifically, in the initial phase, referred to as the *warm-start* round, all clients are required to participate in the local training and the server collects their noisy updates for *indirect* client partitioning (without updating the global model). Then a simple yet powerful clustering method utilizing the Gaussian Mixture Models (GMMs) algorithm is employed, replacing the original *direct* partitioning based on exposed privacy budgets (see Alg. 1).

Privacy Analysis. According to the post-processing immunity property of DP [6], which states that arbitrary data-independent transformations to differentially private outputs will not affect their privacy guarantees, the clustering procedure will not introduce additional privacy costs for each client. Then Algorithm 1 satisfies personalized DP (as stated in Definition 1) [21].

4.2 Experimental Results

We assess the effectiveness of the clustering-based PFA algorithm by evaluating the clustering accuracy after the warm-start round and the test accuracy of the global model achieved at the end of training. We consider three potential multimodal privacy budget distributions (a mixture of two or three different Gaussian distributions, see Table 1)[3]. Due to the space limitation, please refer to [21] for the detailed experimental setup.

[3]This assumption is supported by previous observations which have shown that a bimodal distribution is quite universal in a wide range of complex social systems [31].

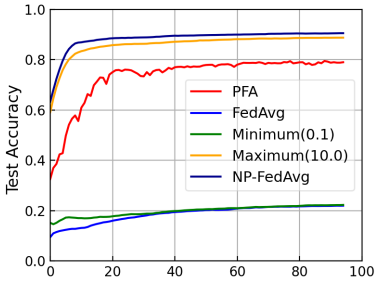


Figure 4: The test accuracy versus communication rounds evaluated on MNIST-LogR in non-IID data setting with privacy budget distribution of MixGauss1.

Effects of the privacy budget distribution. Fig. 3 shows the results of clustering obtained after the warm-start phase in MNIST-LogR experiments, which were conducted with NIID(2) data and varying privacy budget distributions. In all plots, we utilize various markers to represent the predicted cluster index and the real privacy preference of each client. Additionally, we use three different colors to indicate the resulting partitions. Experiment results show an obvious consistency between the L^2-norms clustering and the ground truths (based on clients' real privacy budgets).

Evaluation of the end-to-end PFA framework. In Fig. 4, we report the test accuracy versus communication rounds evaluated on MNIST-LogR in non-IID data setting with privacy budget distribution of MixGauss1. Unlike the results reported by Liu et al. [21], we do not compare the weighted average (WeiAvg) and the communication-efficient version of PFA (i.e., PFA+) here since these two methods are dependent on the values of clients' privacy budgets, which is no longer available in our considered scenario. Just as we expected, the distinct utility advantages of PFA over the baseline methods FedAvg and Minimum remain due to the correct clustering results. Although it has worse accuracy than the non-private baseline (NP-FedAvg), PFA still reaches a reasonable level of model utility, while the FedAvg with PDP becomes ineffective.

5 CONCLUSION AND FUTURE WORK

We propose an effective method for indirect privacy preferences estimation based on L^2-norm clustering in the PDP-FL setting. Then we integrate this approach into the vanilla PFA framework to address potential privacy leakage issues arising from exposed privacy budgets. Our future work will focus on (1) generalizing the clustering strategy to the more challenging cases where clients' privacy budgets are relatively uniform or more difficult to differentiate; (2) conducting extensive empirical evaluations on larger and more diverse datasets for deeper explorations into the effectiveness and scalability of our proposed approach.

6 ACKNOWLEDGEMENT

This research has been funded in part by National Science Foundation of China (NSFC) 62172423 and 62206207, National Science Foundation (NSF) CNS-2124104, CNS-2125530, National Institute of Health (NIH) R01ES033241, R01LM013712. The corresponding author is Xiaofeng Meng.

REFERENCES

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. 2016. Deep learning with differential privacy. In *CCS*.

[2] Naman Agarwal, Peter Kairouz, and Ziyu Liu. 2021. The skellam mechanism for differentially private federated learning. In *NeurIPS*.

[3] Naman Agarwal, Ananda Theertha Suresh, Felix Yu, Sanjiv Kumar, and H. Brendan McMahan. 2018. CpSGD: Communication-Efficient and Differentially-Private Distributed SGD. In *NeurIPS*.

[4] Xiangyi Chen, Steven Z Wu, and Mingyi Hong. 2020. Understanding gradient clipping in private sgd: A geometric perspective. In *NeurIPS*.

[5] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *TCC*.

[6] Cynthia Dwork, Aaron Roth, et al. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science* 9, 3-4 (2014), 211–407.

[7] Robin C Geyer, Tassilo Klein, and Moin Nabi. 2017. Differentially private federated learning: A client level perspective. In *arXiv preprint arXiv:1712.07557*.

[8] Xin Gu, Gautam Kamath, and Zhiwei Steven Wu. 2023. Choosing public datasets for private machine learning via gradient subspace distance. In *arXiv preprint arXiv:2303.01256*.

[9] Filip Hanzely, Slavomír Hanzely, Samuel Horváth, and Peter Richtárik. 2020. Lower bounds and optimal algorithms for personalized federated learning. In *NeurIPS*.

[10] Jonathan Huang. 2005. Maximum likelihood estimation of Dirichlet distribution parameters. *CMU Technique report* 18 (2005).

[11] Zach Jorgensen, Ting Yu, and Graham Cormode. 2015. Conservative or liberal? Personalized differential privacy. In *ICDE*.

[12] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. 2021. Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning* 14, 1–2 (2021), 1–210.

[13] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. 2020. SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. In *ICML*.

[14] Alex Krizhevsky, Geoffrey Hinton, et al. 2009. Learning multiple layers of features from tiny images.

[15] Yann LeCun, Corinna Cortes, and CJ Burges. 2010. MNIST handwritten digit database. *ATT Labs [Online]. Available: http://yann.lecun.com/exdb/mnist* 2.

[16] Qinbin Li, Yiqun Diao, Quan Chen, and Bingsheng He. 2022. Federated learning on non-iid data silos: An experimental study. In *ICDE*.

[17] Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. 2021. Ditto: Fair and robust federated learning through personalization. In *ICML*.

[18] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. 2020. Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems* 2 (2020), 429–450.

[19] Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, and Zhihua Zhang. 2019. On the convergence of fedavg on non-iid data. In *arXiv preprint arXiv:1907.02189*.

[20] Haowen Lin, Jian Lou, Li Xiong, and Cyrus Shahabi. 2021. Semifed: Semi-supervised federated learning with consistency and pseudo-labeling. *arXiv preprint arXiv:2108.09412* (2021).

[21] Junxu Liu, Jian Lou, Li Xiong, Jinfei Liu, and Xiaofeng Meng. 2021. Projected federated averaging with heterogeneous differential privacy. In *VLDB*.

[22] Ken Liu, Shengyuan Hu, Steven Z Wu, and Virginia Smith. 2022. On privacy and personalization in cross-silo federated learning. In *NeurIPS*.

[23] Jian Lou and Yiu-ming Cheung. 2018. Uplink communication efficient differentially private sparse optimization with feature-wise distributed data. In *AAAI*.

[24] Jian Lou and Yiu-ming Cheung. 2020. An uplink communication-efficient approach to featurewise distributed sparse optimization with differential privacy. *IEEE Transactions on Neural Networks and Learning Systems* 32, 10 (2020), 4529–4543.

[25] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *AISTATS*.

[26] Milad Nasr, Reza Shokri, and Amir Houmansadr. 2019. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *S&P*.

[27] Ben Niu, Yahong Chen, Boyang Wang, Zhibo Wang, Fenghua Li, and Jin Cao. 2021. AdaPDP: Adaptive personalized differential privacy. In *INFOCOM*.

[28] Maxence Noble, Aurélien Bellet, and Aymeric Dieuleveut. 2022. Differentially private federated learning on heterogeneous data. In *AISTATS*.

[29] Farnaz Tahmasebian, Jian Lou, and Li Xiong. 2022. Robustfed: a truth inference approach for robust federated learning. In *CIKM*.

[30] Zhibo Wang, Jiahui Hu, Ruizhao Lv, Jian Wei, Qian Wang, Dejun Yang, and Hairong Qi. 2018. Personalized privacy-preserving task allocation for mobile crowdsensing. *IEEE Transactions on Mobile Computing* 18, 6 (2018), 1330–1341.

[31] Ye Wu, Changsong Zhou, Jinghua Xiao, Jürgen Kurths, and Hans Joachim Schellnhuber. 2010. Evidence for a bimodal distribution in human communication. *PNAS* 107, 44 (2010), 18803–18808.

[32] Lei Yu, Ling Liu, Calton Pu, Mehmet Emre Gursoy, and Stacey Truex. 2019. Differentially private model publishing for deep learning. In *S&P*.

[33] Mikhail Yurochkin, Mayank Agarwal, Soumya Ghosh, Kristjan Greenewald, Nghia Hoang, and Yasaman Khazaeni. 2019. Bayesian nonparametric federated learning of neural networks. In *ICML*.

[34] Yingxue Zhou, Steven Wu, and Arindam Banerjee. 2021. Bypassing the Ambient Dimension: Private SGD with Gradient Subspace Identification. In *ICLR*.