西北工業大學

Research Report

得分：

Student ID 2024280038

Name Jason Rich Darmawan

Course Name High Performance Computing

Date 2025/01/07

西北工业大学研究生院

Challenges in Federated Learning for Medical Imaging, Existing Methods, and Their Limitations

Jason Rich Darmawan
School of Computer Science
Northwestern Polytechnical University
Xi'an, China
jasonrichdarmawan@mail.nwpu.edu.cn

Abstract

Objectives: Federated Learning (FL) faces several challenges, including label noise, heterogenous clients' dataset distributions across clients, and vulnerability to adversarial attacks such as byzantine, backdoor, and free-rider attacks. Additionally, FL suffers from issues like incentivizing participation from clients with small and distinct dataset distributions, high communication costs, and fault tolerance. Despite the growing body of research, existing methods often address specific challenges in isolation. This review aims to comprehensively identify and analyze the challenges and properties associated with FL, define the basic concepts of existing method categories, and highlight their limitations.

Methods: A search was conducted using Google Scholar and IEEE Xplore, using the search term "federated learning" to identify relevant studies.

Results: The review references a total of 43 papers, identifies five general challenges, two properties associated with FL, and categorizes nine existing methods, and highlight their limitations.

Keywords—federated learning

摘要

目的：联邦学习 (FL) 面临多项挑战，包括标签噪声、客户端之间的异构数据集分布以及易受拜占庭攻击、后门攻击和搭便车攻击等对抗性攻击。此外，联邦学习还存在一些问题，例如激励数据集分布较小且不同的客户端参与、通信成本高以及容错能力差。尽管研究成果不断增加，但现有方法通常只能孤立地应对特定挑战。本综述旨在全面识别和分析与联邦学习相关的挑战和属性，定义现有方法类别的基本概念，并强调其局限性。

方法：使用 Google Scholar 和 IEEE Xplore 进行搜索，使用搜索词“联邦学习”来查找相关研究。

结果：该评论总共引用了 43 篇论文，确定了五个普遍的挑战、两个与 FL 相关的属性，并对九种现有方法进行了分类，并强调了它们的局限性。

关键词—federated learning

1 Introduction

Federated Learning is an approach to train Deep Learning Models across distributed datasets without requiring participants to share the participant's private raw data. However, its performance can be influenced by various factors, including client's dataset size [1].

Solutions to address client's dataset size problem typically fall into two categories:

1. **Data Augmentation**: techniques that synthetically create new samples [2], [3], [4].
2. **Semi-Supervised Pseudo-Labeling**: methods that train Deep Learning models with partially labeled dataset and use the models to predict the labels for the unlabeled samples [2], [5], [6]

While these approaches offer benefits, they can worsen the problems of noisy labels, impacting the model performance [7]. Furthermore, Federated Learning faces additional challenges, including:

1. **Incentives for Parties with Small and Distinct Dataset:** Sub-optimal performance for parties with small and distinct dataset [1].
2. **Communication Costs**: High communication costs due to communication between clients and the server [8].
3. **Fault Tolerance**: Managing unreliable participants [9].

However, existing methods tend to focus on addressing specific challenges in isolation, leaving gaps in realistic solutions for Federated Learning.

The existing Federated Learning methods [10], [11] are designed to train a global model for different parties without requiring participants to share the participant's private raw data [1]. This approach typically involves two main steps:

1. **Server-Side Collaboration**: A central server aggregates the model parameters from participating clients and distributes the averaged parameters.
2. **Client-Side Optimization**: Clients locally optimize the distributed model using their private data [1].

Several reviews have addressed various aspects of Federated Learning, offering insights into specific challenges and domains:

1. **Federated Learning Categories** [12]: A review categorizing Federated Learning into:
    a. Horizontal Federated Learning: Training across overlapping feature spaces but distinct samples (e.g., collaboration between two banks to train a Federated Learning model for credit risk prediction).
    b. Vertical Federated Learning: Training across distinct feature spaces but overlapping samples (e.g., datasets from a bank and an e-commerce platform).
    c. Federated Transfer Learning: Training on small, shared samples to learn common representations despite clients' datasets have distinct feature spaces and samples (e.g. datasets from a bank in China and an e-commerce platform in the United States)
2. **Data Heterogeneity Problem** [13]: Concluded existing methods that addresses the data heterogeneity problem is using rigid data partitioning strategies which do not cover all typical data heterogeneity cases.
3. **Security and Fairness** [14]: Focuses on existing defenses against malicious attacks and existing methods to incentivize active client participation in Federated Learning.
4. **Specific Domain** [15]: Explore Federated Learning applications in healthcare but is overly focused on the domain, limiting its generalizability to other domains [1].

This study aims to address the following research questions within Federated Learning, particularly in the context of medical imaging:

1. What are the challenges, specific problems and associated properties of Federated Learning?
2. How can Federated Learning be protected against malicious parties?
3. How can parties be incentivized to actively participate in Federated Learning in Medical Imaging?
4. How can selected parameters generalize beyond the training set?

This paper makes the following contributions to the field of Federated, with a focus on medical imaging:

1. Identification of five general challenges and their specific sub-challenges in Federated Learning.
2. Identification of two properties associated with Federated Learning.
3. Categorization and analysis of nine existing FL methods categories.

2 Methods

2.1 Eligibility Criteria

The inclusion criteria for this review are as follows:

1. Articles must be written in English.
2. Eligible sources include conference papers, journal articles, and preprint articles.

2.2 Information Sources

The database used for the literature search are:

1. Google Scholar.
2. IEEE Xplore.

2.3 Search Strategy

The search was conducted using the following keyword: “federated learning”

2.4 Study Selection

The selected studies were categorized based on challenges in Federated Learning, properties of Federated Learning, and methods in Federated Learning.

3 Results

3.1 Overview of Challenges in Federated Learning

3.1.1 Generalization Ability

The ability of Federated Learning models to generalize effectively is hindered by challenges related to heterogeneous clients’ datasets distribution and the presence of noisy labels in the dataset.

General Problem:

1. **Heterogeneous Clients’ Dataset Distribution**: Datasets across clients often vary in distribution [16].
2. **Noisy Labels**: Labels within datasets may be noisy, adversely affecting model performance [17].

Specific Problems:

1. **Different Local Minimum Problem**: Clients optimize local models towards their own dataset's local minimum, leading to differing directions for optimization. This results in extended training time and degraded performance [1].
2. **Unseen Data and Distributions Problems**: Models trained solely on datasets provided by participating clients often perform poorly on unseen data or distributions. Existing methods primarily address unseen dataset distributions problem while neglecting the unseen data problem [18].
3. **Noisy Label Problem**: Variability in label distributions causes aggregated updates from clients to yield high loss when predicting client-specific samples [17]. Traditional Deep Learning models methods such as sample filtering [19], [20], [21] based on loss ranking will have poor performance due to the noisy label problem.
4. **Class Imbalance Problem**: Imbalanced class distributions lead to accuracy degradation [22]. For example, training FedAvg [23] on imbalanced dataset results in a 7.92% accuracy loss compared to balanced dataset [22].

3.1.1.1 Protection Against Malicious Attacks

Federated Learning decentralized nature makes it vulnerable to adversarial actions by malicious parties.

General Problem: A small set of malicious parties can manipulate the training process by uploading poisoned updates to the server [1].

Specific Problems:

1. **Byzantine Attacks**: Malicious parties aim to degrade model convergence and performance [1]. The byzantine attacks can be divided into two categories based on the target to manipulate: a) **Data-Based Attacks** [24], [25]: Poison client training data, b) **Model-Based Attacks** [26]: Manipulate model parameters directly. In addition, the attacks can be divided into two categories based on the updates from the clients: a) **Non-Sybil Attacks**: Each malicious client uploads different updates, complicating detection due to heterogenous distribution between clients' datasets [27], b) **Sybil Attacks** [28]: Malicious clients upload identical updates, exploiting assumptions in defenses based on Euclidean distance [27].
2. **Backdoor Attacks**: Malicious clients introduce backdoor triggers in training data or model parameters to cause targeted misclassifications [1]. The backdoor attacks can be divided into two categories: a) **Centralized Backdoor Attacks** [29]: Each malicious client injects the same trigger during training, b) **Distributed Backdoor Attacks** [30]: Each malicious client injects only parts of the trigger. Although each client only injects a part of the trigger, when the server aggregates the updates from

clients, the trigger parts are also aggregated. The following are properties of specific backdoor attacks:

a. A distributed backdoor attack proposed by [30] divides the backdoor trigger across malicious clients may cause small decrease for the main task accuracy right after the attack. However, the Federated Learning model performance will get back to normal after a few rounds of training.
b. A backdoor attack proposed by [31] uses the same backdoor trigger across malicious clients, resulting in similar updates across malicious clients and different from benign clients [1].
c. A backdoor attack proposed by [32] specifically targets model parameters that are less likely to undergo significant updates during the training process. This strategy is designed to increase the survival rate of the backdoor trigger, ensuring its persistence and effectiveness even after multiple rounds of training and updates without the attack.
d. A backdoor attack proposed by [33] minimizes the distance between the backdoored model and the global model parameters to bypass the anomaly detection-based defenses.
e. A backdoor attack proposed by [34] employs adversarial training to adapt the backdoor trigger. This approach aims to make the trigger more resistant to defenses by creating a backdoor trigger that is difficult for the global model to unlearn, thereby ensuring its continued effectiveness.

3. **Free Rider Attack**: Malicious clients upload fake updates to access global models without contributing to training [1].

3.1.2 Incentivizing Participation

Federated Learning lacks mechanisms to encourage active participation from all parties.

General Problem: Parties require incentives, such as improved model accuracy for their data, to participate actively [35].

Specific Problems: **Sub-Optimal Performance for Small Datasets**: Clients with small or distinct dataset distribution often experience sub-optimal performance [1].

3.1.3 Communication Costs

Federated Learning depends on large-scale distributed training, where network bandwidth becomes a limiting factor [8].

Challenge: 99.9% of the gradients uploaded at each communication round was found redundant [8]. To address this problem, [8] proposed to compress the gradients.

3.1.4 Fault Tolerance

Federated Learning systems must handle failures in communication and computation during training.

Challenges:

1. **Communication Problem**: Clients may disconnect temporarily or drop out entirely [9].
2. **Computation Problem**: Clients may require varying amounts of time to complete training and send updates. Systems must remain operational without being delayed by slower clients [9].

3.2 Overview of Properties in Federated Learning

Federated Learning exhibits certain properties that influence its behavior and the design of defense mechanisms:

1. **Euclidean Distance**: The Euclidean distance between two benign clients is as expected low [27].
2. **Cosine Similarity**: The cosine similarity between the gradients of two benign clients is unexpectedly low, and irrespective of whether the clients' datasets have heterogonous or homogonous distribution [27]. This phenomenon occurs because, as the global model converges, the gradients of client's model approach near-zero values. Consequently, the cosine similarity between two clients' gradients approaches zero.

3.3 Overview of Existing Methods in Federated Learning

This section provides a structured overview of existing methods addressing various challenges in Federated Learning

3.3.1 Methods Addressing the Different Local Minimum Problem

1. **Penalizing Local Directions**: Methods like [36] penalize local directions using the global model output on local private data. However, clients must download and run the global model, increasing communication and computation costs as model parameters scale.
2. **Global Statistics-Based Loss Functions**: Methods like [37] utilize local statistics uploaded by clients to compute global statistics at the server, which then constructs a loss function. However, clients must download the updated global model, increasing communication costs as model parameters scale.

3. **Conditional Generative Adversarial Networks (cGANs)**: Methods like [38] employ a classification network and cGANs, requiring clients to exchange their classifier and generator models. The need to upload and download these components increases communication costs as the model parameters scale.
4. **Model Output Exchange and Shared Unlabeled Dataset**: Methods like [39] propose sharing an unlabeled dataset among clients, using a teacher network for labeling and exchanging model outputs instead of parameters. [39] identifies two challenges associated with model output exchange:
    a. **Global Logit Construction**: Aggregated local logits resemble local datasets, making retraining nearly equivalent to training directly on the local data without distillation [39]. To address this challenge, [39] propose using a teacher network to label the shared dataset and calculate local logits. The labeled data is then combined for retraining [39].
    b. **Heterogonous Distributions**: Logits from heterogonous distributions correspond to different local minima, slowing convergence [39]. To address this challenge, [39] propose reducing the entropy of global logits.
5. **Data Augmentation with Mediators**: Methods like [22] use data augmentation and mediators to balance class distribution and aggregate local model parameters. The following are the proposed steps:
    a. The server computes the global class distribution based on aggregated local class distributions [22].
    b. Clients perform data augmentation using the global class distribution [22].
    c. Mediators select clients to form a uniform class distribution and aggregate their local model parameters [22].
    d. Aggregated parameters are sent by the mediators to the server [22].
6. **Reinforcement Learning for Client Selection**: Methods like [40] use reinforcement learning, where an agent interacts with the environment and receives rewards to select clients that can enhance global model accuracy.

3.3.2 Methods Addressing Malicious Attacks

1. **Similarity-Based Detection**: Methods like [41] use ReLU-clipped cosine similarity to identify malicious clients by comparing updates with benign clients. However, the server may still select malicious clients in subsequent rounds.
2. **Client Selection Based on Past Accuracy**: Methods like [27] addresses the above limitation by choosing clients based on their past accuracies. In addition, [27] propose the following to address sybil attacks [28] and non-sybil attacks.
    a. **Sybil attacks** [28]: [27] use cosine similarity to detect sybil attacks, where malicious clients upload identical updates.

b. **Non-Sybil Attacks**: These pose a greater challenge due to heterogenous distributions between clients' datasets. To mitigate this, [27] reduces the variance between local updates.

3.3.3 Methods to Incentivize Active Participation

1. **Client Contribution Evaluation**: Methods like [1] evaluate the specific client contribution based on the difference in the global model accuracy when the specific client update is excluded.
2. **Mixture of Client Distributions:** Methods like [42] optimizes the global model for any target distribution based on a mixture of client distributions instead of a specific distribution.
3. **Dual-loss Approach**: Methods like [43] divides the loss function into two: a) a local task loss, and b) a consistency across client loss. This method allows each client to build a custom model architecture while still benefitting from the learnings of other clients [43]. However, this approach risks overfitting on small datasets [17].

4 Conclusion

Our review indicates that current research in Federated Learning predominantly focuses on addressing individual challenges in isolation, often overlooking the need for realistic solutions that can address multiple issues simultaneously. A significant portion of the research has been concentrated on dealing with heterogeneity in clients' datasets distributions. While this is crucial, this narrow focus may limit the applicability and effectiveness of Federated Learning in real-world scenarios, where multiple factors—such as communication costs, security threats, and incentive mechanisms must be considered together.

Notable advancements, such as the work by [39], which proposes training Federated Learning models by exchanging model outputs instead of parameters, and the findings by [8], showing that 99.9% of gradient exchanges are redundant. These insights suggest that reducing communication overhead by focusing on model outputs could be a viable solution, potentially alleviating some of the scalability issues associated with Federated Learning.

Based on our analysis, we recommend exploring methods that integrate model output exchange as a primary strategy while also addressing other core challenges, such as dataset distribution heterogeneity, malicious attacks, and incentive mechanisms. A more unified approach that considers these challenges together could pave the way for more practical and efficient Federated Learning systems.

5 Acknowledgement

Thanks Associate Professor Huanjie Tao for his invaluable guidance.

6 References

[1] W. Huang *et al.*, "Federated Learning for Generalization, Robustness, Fairness: A Survey and Benchmark," *IEEE Trans Pattern Anal Mach Intell*, vol. 46, no. 12, pp. 9387–9406, 2024, doi: 10.1109/TPAMI.2024.3418862.

[2] K. Nishi, Y. Ding, A. Rich, and T. Hollerer, "Augmentation Strategies for Learning With Noisy Labels," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 8022–8031.

[3] E. D. Cubuk, B. Zoph, D. Mané, V. Vasudevan, and Q. V Le, "AutoAugment: Learning Augmentation Strategies From Data," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 113–123. doi: 10.1109/CVPR.2019.00020.

[4] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V Le, "Randaugment: Practical automated data augmentation with a reduced search space," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 3008–3017. doi: 10.1109/CVPRW50498.2020.00359.

[5] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," in *International Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=r1Ddp1-Rb

[6] K. Sohn *et al.*, "FixMatch: simplifying semi-supervised learning with consistency and confidence," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, in NIPS '20. Red Hook, NY, USA: Curran Associates Inc., 2020.

[7] Y. Wei, Y. Deng, C. Sun, M. Lin, H. Jiang, and Y. Peng, "Deep learning with noisy labels in medical prediction problems: a scoping review," *Journal of the American Medical Informatics Association*, vol. 31, no. 7, pp. 1596–1607, Nov. 2024, doi: 10.1093/jamia/ocae108.

[8] Y. Lin, S. Han, H. Mao, Y. Wang, and B. Dally, "Deep Gradient Compression: Reducing the Communication Bandwidth for Distributed Training," in *International Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=SkhQHMW0W

[9] V. Smith, C.-K. Chiang, M. Sanjabi, and A. Talwalkar, “Federated multi-task learning,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in NIPS’17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 4427–4437.

[10] A. Hard *et al.*, “Federated Learning for Mobile Keyboard Prediction,” 2018. [Online]. Available: https://arxiv.org/abs/1811.03604

[11] S. Pati *et al.*, “Federated learning enables big data for rare cancer boundary detection,” *Nat Commun*, vol. 13, no. 1, p. 7346, 2022, doi: 10.1038/s41467-022-33407-5.

[12] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated Machine Learning: Concept and Applications,” *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, Jan. 2019, doi: 10.1145/3298981.

[13] Q. Li, Y. Diao, Q. Chen, and B. He, “Federated Learning on Non-IID Data Silos: An Experimental Study,” in *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, 2022, pp. 965–978. doi: 10.1109/ICDE53745.2022.00077.

[14] Q. Li *et al.*, “A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection,” *IEEE Trans Knowl Data Eng*, vol. 35, no. 4, pp. 3347–3366, 2023, doi: 10.1109/TKDE.2021.3124599.

[15] D. C. Nguyen *et al.*, “Federated Learning for Smart Healthcare: A Survey,” *ACM Comput. Surv.*, vol. 55, no. 3, Feb. 2022, doi: 10.1145/3501296.

[16] N. Shoham *et al.*, “Overcoming Forgetting in Federated Learning on Non-IID Data,” 2019. [Online]. Available: https://arxiv.org/abs/1910.07796

[17] X. Ji *et al.*, “FedFixer: mitigating heterogeneous label noise in federated learning,” in *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, in AAAI’24/IAAI’24/EAAI’24. AAAI Press, 2025. doi: 10.1609/aaai.v38i11.29179.

[18] H. Yuan, W. R. Morningstar, L. Ning, and K. Singhal, “What Do We Mean by Generalization in Federated Learning?,” in *International Conference on Learning Representations*, 2022. [Online]. Available: https://openreview.net/forum?id=VimqQq-i_Q

[19] B. Han *et al.*, “Co-teaching: robust training of deep neural networks with extremely noisy labels,” in *Proceedings of the 32nd International Conference on Neural*

Information Processing Systems, in NIPS'18. Red Hook, NY, USA: Curran Associates Inc., 2018, pp. 8536–8546.

[20] X. Yu, B. Han, J. Yao, G. Niu, I. Tsang, and M. Sugiyama, "How does Disagreement Help Generalization against Label Corruption?," in *Proceedings of the 36th International Conference on Machine Learning*, K. Chaudhuri and R. Salakhutdinov, Eds., in Proceedings of Machine Learning Research, vol. 97. PMLR, Oct. 2019, pp. 7164–7173. [Online]. Available: https://proceedings.mlr.press/v97/yu19b.html

[21] L. Ju *et al.*, "Improving Medical Images Classification With Label Noise Using Dual-Uncertainty Estimation," *IEEE Trans Med Imaging*, vol. 41, no. 6, pp. 1533–1546, 2022, doi: 10.1109/TMI.2022.3141425.

[22] M. Duan *et al.*, "Astraea: Self-Balancing Federated Learning for Improving Classification Accuracy of Mobile Deep Learning Applications," in *2019 IEEE 37th International Conference on Computer Design (ICCD)*, 2019, pp. 246–254. doi: 10.1109/ICCD46524.2019.00038.

[23] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, A. Singh and J. Zhu, Eds., in Proceedings of Machine Learning Research, vol. 54. PMLR, Jan. 2017, pp. 1273–1282. [Online]. Available: https://proceedings.mlr.press/v54/mcmahan17a.html

[24] B. van Rooyen, A. Menon, and R. C. Williamson, "Learning with Symmetric Label Noise: The Importance of Being Unhinged," in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., Curran Associates, Inc., 2015. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2015/file/45c48cce2e2d7fbdea1afc51c7c6ad26-Paper.pdf

[25] J. Zhang *et al.*, "BadLabel: A Robust Perspective on Evaluating and Enhancing Label-Noise Learning," *IEEE Trans Pattern Anal Mach Intell*, vol. 46, no. 6, pp. 4398–4409, Jun. 2024, doi: 10.1109/TPAMI.2024.3355425.

[26] M. Baruch, G. Baruch, and Y. Goldberg, "A little is enough: circumventing defenses for distributed learning," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Red Hook, NY, USA: Curran Associates Inc., 2019.

[27] W. Wan, S. Hu, jianrong Lu, L. Y. Zhang, H. Jin, and Y. He, “Shielding Federated Learning: Robust Aggregation with Adaptive Client Selection,” in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, L. De Raedt, Ed., International Joint Conferences on Artificial Intelligence Organization, Jan. 2022, pp. 753–760. doi: 10.24963/ijcai.2022/106.

[28] C. Fung, C. J. M. Yoon, and I. Beschastnikh, “The Limitations of Federated Learning in Sybil Settings,” in *23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*, San Sebastian: USENIX Association, Oct. 2020, pp. 301–316. [Online]. Available: https://www.usenix.org/conference/raid2020/presentation/fung

[29] H. Zhong, C. Liao, A. C. Squicciarini, S. Zhu, and D. Miller, “Backdoor Embedding in Convolutional Neural Network Models via Invisible Perturbation,” in *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy*, in CODASPY ’20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 97–108. doi: 10.1145/3374664.3375751.

[30] C. Xie, K. Huang, P.-Y. Chen, and B. Li, “DBA: Distributed Backdoor Attacks against Federated Learning,” in *International Conference on Learning Representations*, 2020. [Online]. Available: https://openreview.net/forum?id=rkgyS0VFvr

[31] H. Wang *et al.*, “Attack of the tails: yes, you really can backdoor federated learning,” in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, in NIPS ’20. Red Hook, NY, USA: Curran Associates Inc., 2020.

[32] Z. Zhang *et al.*, “Neurotoxin: Durable Backdoors in Federated Learning,” in *Proceedings of the 39th International Conference on Machine Learning*, K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, Eds., in Proceedings of Machine Learning Research, vol. 162. PMLR, Jan. 2022, pp. 26429–26446. [Online]. Available: https://proceedings.mlr.press/v162/zhang22w.html

[33] P. Fang and J. Chen, “On the Vulnerability of Backdoor Defenses for Federated Learning,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 10, pp. 11800–11808, Jun. 2023, doi: 10.1609/aaai.v37i10.26393.

[34] H. Zhang, J. Jia, J. Chen, L. Lin, and D. Wu, “A3FL: Adversarially Adaptive Backdoor Attacks to Federated Learning,” in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., Curran Associates, Inc., 2023, pp. 61213–61233. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/c07d71ff0bc042e4b9acd626a79597fa-Paper-Conference.pdf

[35] T. Song, Y. Tong, and S. Wei, “Profit Allocation for Federated Learning,” in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 2577–2586. doi: 10.1109/BigData47090.2019.9006327.

[36] Q. Li, B. He, and D. Song, “Model-Contrastive Federated Learning,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 10713–10722.

[37] W. Huang, M. Ye, Z. Shi, H. Li, and B. Du, “Rethinking Federated Learning with Domain Shift: A Prototype View,” in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 16312–16322. doi: 10.1109/CVPR52729.2023.01565.

[38] Y. Wu *et al.*, “FedCG: Leverage Conditional GAN for Protecting Privacy and Maintaining Competitive Performance in Federated Learning,” in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, L. De Raedt, Ed., International Joint Conferences on Artificial Intelligence Organization, Jan. 2022, pp. 2334–2340. doi: 10.24963/ijcai.2022/324.

[39] S. Itahara, T. Nishio, Y. Koda, M. Morikura, and K. Yamamoto, “Distillation-Based Semi-Supervised Federated Learning for Communication-Efficient Collaborative Training With Non-IID Private Data,” *IEEE Trans Mob Comput*, vol. 22, no. 1, pp. 191–205, 2023, doi: 10.1109/TMC.2021.3070013.

[40] H. Wang, Z. Kaplan, D. Niu, and B. Li, “Optimizing Federated Learning on Non-IID Data with Reinforcement Learning,” in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, 2020, pp. 1698–1707. doi: 10.1109/INFOCOM41043.2020.9155494.

[41] X. Cao, M. Fang, J. Liu, and N. Z. Gong, “FLTrust: Byzantine-robust Federated Learning via Trust Bootstrapping,” in *Proceedings 2021 Network and Distributed System Security Symposium*, Reston, VA: Internet Society, 2021. doi: 10.14722/ndss.2021.24434.

[42] M. Mohri, G. Sivek, and A. T. Suresh, “Agnostic Federated Learning,” in *Proceedings of the 36th International Conference on Machine Learning*, K. Chaudhuri and R. Salakhutdinov, Eds., in Proceedings of Machine Learning Research, vol. 97. PMLR, Jan. 2019, pp. 4615–4625. [Online]. Available: https://proceedings.mlr.press/v97/mohri19a.html

[43] D. Makhija, X. Han, N. Ho, and J. Ghosh, “Architecture Agnostic Federated Learning for Neural Networks,” in *Proceedings of the 39th International Conference on*

Machine Learning, K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, Eds., in Proceedings of Machine Learning Research, vol. 162. PMLR, Jan. 2022, pp. 14860–14870. [Online]. Available: https://proceedings.mlr.press/v162/makhija22a.html