Improved Hybrid Approach Combining Differential Privacy and Byzantine Fault Tolerance for Addressing Poisoning Attacks in Federated Learning

Yakubu Galadima Ibrahim
Dept. of Computer Science Al-Qalam University Katsina. Katsina,Nigeria
yakubugaladima01@gmail.com

Armaya'u Zango Umar
Dept. Soft. Eng. and Cybersec. Al-Qalam University Katsina. Katsina, Nigeria.
azumar@auk.edu.ng

Umar Abdulmalik
Dept. of Computer Science. Al-Qalam University Katsina. Katsina,Nigeria
abdulmalikumar1997@gmail.com

***Abstract*—Federated Learning (FL) offers a privacy-aware alternative to centralized machine learning by enabling decentralized model training without raw data sharing. However, its susceptibility to adversarial threats, particularly poisoning attacks, and the risk of privacy leakage through model updates remain critical challenges. To address these issues, this research proposes an improved hybrid defense mechanism that enhances the Differentially Private Byzantine- Resilient Momentum (DP-BREM) framework through the integration of Adaptive Differential Privacy (ADP). The proposed model introduces a dynamic noise scaling strategy that adjusts privacy noise based on training round and model performance, enabling a more efficient privacy-utility tradeoff. Additionally, a Byzantine-resilient aggregation method is incorporated to mitigate the impact of malicious client updates. The framework was implemented and evaluated using the MNIST dataset under simulated adversarial conditions involving 20% Byzantine clients. Results demonstrate that the Adaptive DP-BREM significantly outperforms the baseline DP-BREM with fixed noise, achieving higher classification accuracy at lower privacy budgets, reducing cumulative privacy loss by approximately 1.3 ε, and enhancing adversarial detection rates across varying attack strengths. These findings underscore the effectiveness of adaptive privacy mechanisms in preserving model utility and robustness, thereby advancing the reliability of federated learning in privacy-sensitive and adversarial environments.**

***Keywords*—Federated learning, Adaptive Differential Privacy, Byzantine Fault Tolerance**

I. INTRODUCTION

Traditional machine learning (ML) systems rely heavily on centralized data collection, where raw data from various sources is aggregated into a single repository for model training. While this paradigm has driven significant advances on in domains such as computer vision and natural language processing, it introduces critical privacy and security concerns, particularly in sensitive sectors like healthcare, finance, and national security. Additionally, centralized systems encounter scalability bottlenecks, including high communication costs and computational demands (McMahan et al., 2017). Federated Learning (FL) has emerged as a promising alternative to centralized learning by enabling decentralized training of machine learning models. In FL, clients (e.g., mobile devices, hospitals, banks) retain their data locally and exchange only model updates (e.g., gradients or weights) with a central server. This privacy-aware framework allows collaborative model building while preserving data locality (Kairouz et al., 2021). Despite its privacy-centric design, FL is vulnerable to sophisticated adversarial threats. Among these, poisoning attacks are particularly concerning, as malicious clients can submit manipulated updates to degrade or backdoor the global model (Bagdasaryan et al., 2020). Furthermore, recent studies have shown that model updates can inadvertently leak private information, even when raw data remains local (Hitaj et al., 2017). These challenges necessitate defenses that address both robustness and privacy in federated settings. Byzantine Fault Tolerance (BFT) mechanisms, such as the Krum algorithm, are widely adopted to mitigate the effect of malicious or faulty clients. Krum enhances robustness by selecting updates that are statistically closest to the majority, thus excluding potential outliers introduced by adversaries (Blanchard et al., 2017). Differential Privacy (DP) provides formal privacy guarantees by injecting noise into model updates, thereby obscuring any single client's contribution. While DP has been successfully applied in FL to preserve privacy (Abadi et al., 2016), the injection of static noise often introduces significant utility degradation, especially when combined with adversarial threats.The DP-BREM framework (Gu et al., 2023). was recently introduced as a hybrid defense combining DP with robust momentum-based aggregation. It applies DP noise over aggregated momentum vectors to mitigate both privacy leakage and poisoning effects. However, the framework employs a fixed noise parameter, which can lead to excessive privacy loss or utility degradation depending on the learning phase.

To address these limitations, this study proposes an improved hybrid framework that enhances DP-BREM by integrating Adaptive Differential Privacy (ADP). ADP introduces dynamic noise scaling based on training round and model performance, allowing the system to balance privacy and utility more effectively across different stages of learning. This integration aims to preserve the privacy guarantees of DP while strengthening model robustness against poisoning attacks through performance-aware defense adjustments. Thus, the following are the specific objectives of the study:

- To design and implement a federated learning framework that enhances DP-BREM using adaptive noise scaling.
- To evaluate the effectiveness of the improved framework against poisoning attacks using classification accuracy and privacy loss as performance metrics.
- To compare the proposed model with baseline DP-BREM under identical adversarial scenarios.

II. Related Works

Bagdasaryan et al. (2020) conducted an empirical study on backdoor attacks in federated learning, showing how adversaries can subtly manipulate local updates to introduce hidden triggers into the global model. Their findings underscore the urgent need for robust mechanisms that can detect and neutralize these sophisticated attacks. Similarly, Cao et al. (2021) proposed FLTrust, which leverages a small, trusted dataset on the server side to validate incoming gradients. While this enhances Byzantine resilience, it assumes the availability of a reliable reference, which is often impractical in privacysensitive settings. Hard et al. (2018) demonstrated one of the first real-world deployments of federated learning in Google's Gboard application. This case study emphasized the practical utility of FL but also highlighted its challenges, especially in ensuring consistent performance across a large, heterogeneous user base. Kairouz et al. (2019) offered a comprehensive survey of advances and open challenges in FL, identifying crucial gaps in security and privacy that motivate our work. Fung et al. (2020) developed FoolsGold, a defense against Sybil attacks in federated learning. This method identifies groups of malicious clients by analyzing update similarity, but it lacks built-in privacy guarantees. Li, T. et al. (2020) presented a thorough review of federated learning challenges, stressing the importance of secure and robust model training under adversarial threats. Alistarh et al. (2018) and Yin et al. (2018) focused on enhancing Byzantine robustness through techniques like robust stochastic gradient descent and trimmed mean aggregation, respectively. Chen et al. (2017) explored the effect of adversarial clients on distributed learning systems and emphasized the importance of filtering poisoned gradients. Mohri et al. (2019) addressed fairness and robustness through agnostic federated learning, suggesting that defenses must also consider the diverse performance needs of clients. Geyer et al. (2017) proposed client-level DP, optimizing the balance between model utility and privacy by managing local privacy budgets. Li, X. et al. (2021) introduced FedBN, a technique to improve federated learning performance in non-IID environments by maintaining personalized batch normalization layers. This approach enhances convergence under heterogeneous data but does not explicitly address security or privacy. Xie et al. (2019) and Wang et al. (2020) delved into advanced backdoor attack strategies, including distributed and class-targeted methods. Their findings make it clear that static defense mechanisms are inadequate and must be enhanced with adaptive strategies to counter evolving threats.

Prior studies in federated learning have addressed privacy preservation and robustness independently, with comparatively fewer efforts exploring their joint optimization. Abadi et al. (2016) pioneered the use of Differential Privacy (DP) in deep learning through the development of the DP-SGD algorithm, which introduced calibrated noise to gradient updates during training. This method laid the foundation for privacypreserving machine learning but did not consider adversarial threats common in distributed environments. In contrast, Blanchard et al. (2017) proposed the Krum algorithm, a Byzantine-resilient aggregation method designed to identify and filter out malicious updates in distributed learning. Krum selects the update closest to the majority of others based on pairwise Euclidean distances, effectively excluding outliers. While Krum addresses robustness, it does not provide any formal privacy guarantees, making it insufficient for applications involving sensitive data.

To bridge this gap, Gu et al. (2023) introduced DP-BREM (Differentially Private Byzantine-Resilient Momentum Aggregation), which integrates DP with momentum-based aggregation in FL. Instead of applying noise to individual client updates, DP-BREM injects noise into the aggregated momentum, enhancing robustness against poisoning while preserving gradient geometry. However, the method relies on a fixed noise multiplier throughout training, which fails to account for changes in model confidence or convergence dynamics. Building upon this, Zhu and Ling (2022) proposed decoupling the DP mechanism from the robust aggregation process to prevent privacy degradation from compromising defense capabilities. Similarly, Zhang and Hu (2023) explored variance reduction techniques to stabilize differentially private updates under adversarial conditions, though their work did not explicitly integrate momentum-based methods or dynamic privacy adjustment.

Other notable contributions include FLTrust (Cao et al., 2021), which uses a small trusted dataset at the server to guide gradient validation, and FoolsGold (Fung et al., 2020), which detects Sybil-based collusion attacks via similarity analysis. While these methods enhance robustness, they do not incorporate privacy-preserving mechanisms like DP. Shokri et al. (2017) also demonstrated that even partial model sharing can lead to membership inference attacks, reinforcing the need for integrated privacy defenses.

These efforts collectively highlight the need for a unified framework that balances both privacy and robustness. Our research addresses this gap by enhancing DP-BREM with

Adaptive Differential Privacy. Unlike prior models, our approach dynamically scales the noise based on training round and model performance, ensuring tighter privacy guarantees while minimizing the utility loss common in static DP models.

III. METHODOLOGY

This section outlines the methodology employed to develop, implement, and evaluate an improved federated learning defense mechanism. The core innovation lies in refining the existing DP-BREM framework by integrating Adaptive Differential Privacy (ADP), enabling dynamic noise control during training to enhance both privacy and robustness. The proposed method is tested within a simulated federated learning (FL) environment subjected to poisoning attacks. Key components of the methodology include data preparation, system architecture, adversarial threat modeling, defense implementation, and performance evaluation using model accuracy and cumulative privacy loss.

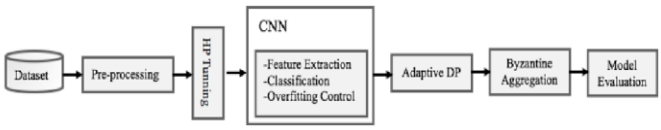

Figure 1: *Improved DP-BREM Methodology*

A. Dataset and Preprocessing

The implementation utilizes the MNIST dataset, a widely adopted benchmark in image classification tasks. MNIST consists of 70,000 grayscale images of handwritten digits (0–9), with 60,000 used for training and 10,000 for testing. The dataset is partitioned across 100 clients to simulate a federated setting, each receiving an independent and identically distributed (IID) subset of data. Each client trains locally on its assigned data using a convolutional neural network (CNN). The global model is updated through iterative communication rounds in which a subset of clients share their locally trained updates with a central server. The model is implemented using PyTorch with the Opacus library for differential privacy accounting. This environment enables evaluation of the proposed method under controlled conditions.

.

B. Model Architecture

The local model architecture comprises a CNN optimized for handwritten digit recognition. It includes two convolutional layers with ReLU activations and max pooling, followed by two

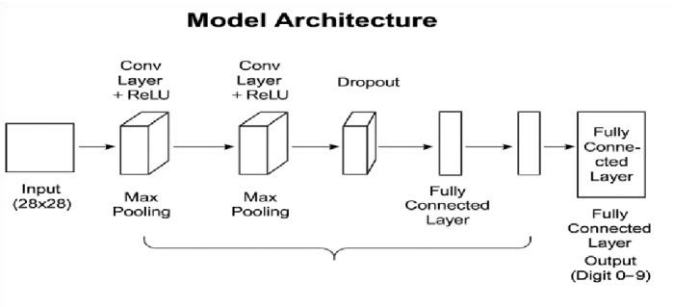

Figure 2. Model Architecture

fully connected layers. Dropout is used to mitigate overfitting. This architecture balances computational efficiency with learning capacity, making it well-suited for FL tasks involving edge devices.

The core contribution of this study is the integration of Adaptive Differential Privacy into the DP-BREM framework. Traditional DP mechanisms use a fixed noise multiplier throughout training, which can either under-protect early rounds or over-noise later stages, degrading model utility. In this improved method, the Gaussian noise added to aggregated momentum updates is dynamically scaled based on two criteria:

Training Round : Early rounds receive higher noise, while later rounds benefit from reduced perturbation. **Model Performance:** Noise is scaled adaptively depending on the model's test accuracy. Improved accuracy permits less noise, while stagnation or degradation triggers increased noise injection.

Mathematically, the noise standard deviation σ for round t is computed as:

$\sigma_t = \max(\sigma \min, \sigma \text{ base} \cdot \gamma^t) \cdot \kappa(\text{Acc}_{t-1})$

where:

γ is the exponential decay rate,

κ is a scaling function based on previous accuracy

σ min , σ base are the base and minimum noise thresholds.

This dynamic strategy ensures optimal trade-off between privacy and utility over the training lifecycle.

C. Algorithm

Algorithm 1: FL-ADPR (Server Orchestration)

This algorithm manages the federated learning cycle, incorporating robust aggregation and privacy accounting.

Input: Global Weights w0; Hyperparameter Set H; Local Datasets {Dk}

Output: Final Global Weights wT; Final Privacy Budget ϵT

1. Initialize w0, m0←0, best_acc←0
2. Initialize P←RDPAccountant(H)
3. **For** round r=1 **to** NUM_ROUNDS **do**
4. Sr←Sample({1,…,N},PARTICIPATION_RATE) (Client Selection)
5. Br,Hr←Designate(Sr,BYZANTINE_FRAC)
6. Collect Updates {Δwk}k∈Sr in parallel from Procedure 2
7. σr←ADAPTIVE_NOISE_SCALING(r,best_acc,H)
8. Δwagg,mr←ROBUST_DP_AGGREGATION({Δwk},mr−1,σr)
9. wr←wr−1+(BASE_LR·LR_DECAYr−1)·Δwagg (Model Update)
10. P←P+RDP_Update(σr) (Privacy Accounting)
11. accr←Evaluate(wr,Dtest)
12. Update best_acc. **If** STAGNATION_CHECK(accr,PATIENCE) **then** RESTART()
13. **End For**
14. ϵT←P.GetPrivacyBudget(DELTA)

Procedure 2: CLIENT_UPDATE

This runs locally on client k.

Input: wglobal,mglobal,Dk,is_byzantine

Output: Δwk

1. **If** is_byzantine **then**
2. Δwk←Generate_Adversarial_Update(r)
3. **Else**
4. g←∇L(wglobal,Dk)
5. gclip←g/max(1,‖g‖2/CLIP_BOUND)
6. Δwk←gclip+mglobal
7. **End If**
8. **Return** Δwk

Procedure 3: ROBUST_DP_AGGREGATION

This runs on the server to ensure security and privacy.

Input: {Δwk}k∈Sr,mprev,σr

Output: Δwdp,mcurr

1. ΔWfiltered←Filter_Extreme_Updates({Δwk})
2. Δwmean←Trimmed_Mean(ΔWfiltered,0.1)
3. Z∼N(0,σr2·I)
4. Δwdp←Δwmean+Z
5. mcurr←Update_Momentum(mprev,Δwdp)
6. **Return** Δwdp,mcurr

IV. RESULTS AND DISCUSSION

In this section we present the experimental results and a comprehensive analysis of the improved federated learning framework integrating Adaptive Differential Privacy (ADP) with Byzantine Fault Tolerance (BFT). The performance of the proposed Adaptive DP-BREM model is evaluated against the baseline DP-BREM model with fixed noise. The evaluation focuses on three core aspects: privacy-accuracy tradeoff, privacy budget consumption, and adversarial detection rate under varying attack strengths.

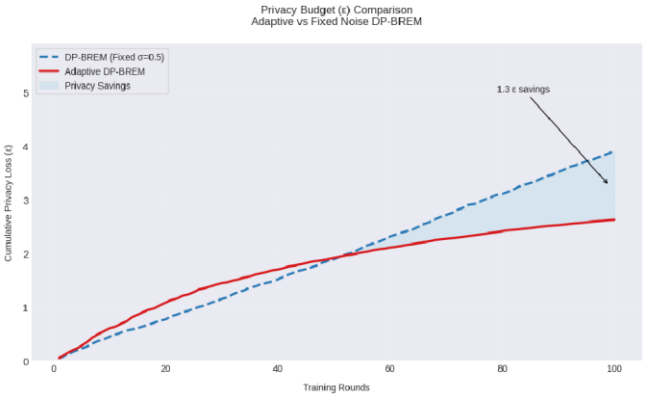

Figure :4 *Privacy Budget Comparison*

To quantify the efficiency of privacy usage over the course of training, Figure 4 presents the cumulative privacy loss (ε) across 100 training rounds for both models. The baseline DPBREM, employing a fixed noise level (σ = 0.5), demonstrates a linear increase in cumulative ε, culminating at approximately ε = 3.9. In contrast, the Adaptive DP-BREM consistently applies noise based on model performance, resulting in a significantly lower cumulative privacy loss of ε = 2.6. The shaded region in the plot denotes the overall privacy savings of 1.3 ε attributable to the adaptive method. This finding highlights the ability of the proposed model to dynamically conserve privacy budget without compromising model performance, making it more suitable for long-running or iterative federated learning deployments. Moreover, this savings aligns with the theoretical advantage of adaptive mechanisms in providing privacy efficiency, a desirable property especially in real-world applications where privacy budgets are finite and costly.

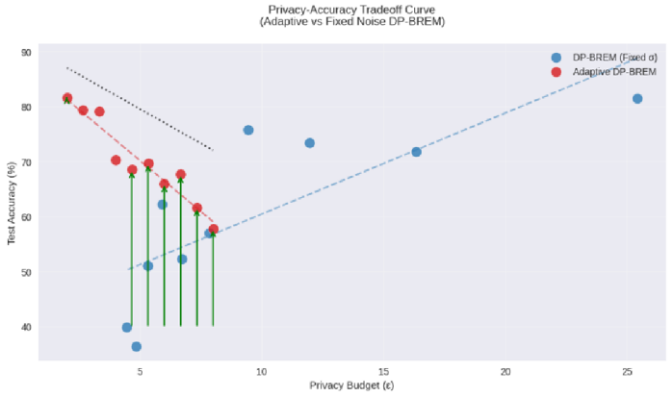

Figure:5 *Privacy Accuracy Tradeoff*

Figure 5 illustrates the privacy-accuracy tradeoff curves for both Adaptive DP-BREM and the standard DP-BREM with fixed Gaussian noise (σ). The horizontal axis represents the privacy budget (ε), while the vertical axis shows the corresponding test accuracy achieved under each privacy constraint. The results demonstrate that the proposed Adaptive DP-BREM significantly outperforms the fixed-noise baseline in maintaining higher accuracy at lower ε values. For instance, while the fixed σ model exhibits sharp declines in test accuracy as the privacy budget tightens, the adaptive approach sustains performance more effectively. This is particularly evident in the privacy budget range of ε = 3 to ε = 8, where the adaptive model maintains a 10–20% improvement in accuracy. The green arrows embedded in the plot depict the gain in accuracy achieved at equivalent privacy levels, thus empirically validating the utility-preserving advantage of dynamic noise scaling. This result underscores a critical contribution of this work: the dynamic adjustment of noise magnitude enables better calibration between privacy guarantees and model utility, particularly under strict privacy requirements. Consequently, the adaptive mechanism mitigates the performance degradation traditionally associated with strong differential privacy enforcement

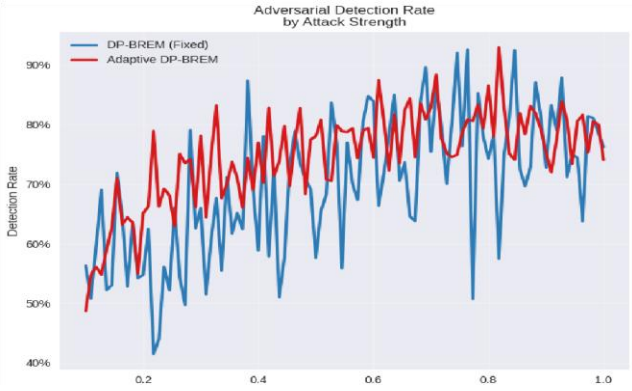

Figure: 6 *Adversarial Detection Rate*

Figure 6 evaluates the robustness of both models against adversarial manipulation through the lens of detection rates across varying attack strengths. The vertical axis represents the detection rate, while the horizontal axis shows the attack strength (ranging from 0.05 to 1.0). Detection was measured based on the system's ability to identify and exclude malicious gradients during federated aggregation. The Adaptive DPBREM consistently achieves higher and more stable detection rates compared to the fixed-noise variant, particularly in the moderate to high attack strength range. While both models exhibit fluctuations due to the stochastic nature of adversarial attacks, the adaptive model maintains a smoother and more resilient profile, often exceeding 80% detection accuracy under intense attacks. In contrast, the fixed σ model exhibits high variance and notable drops in performance at various attack intervals.

This result confirms that adaptive privacy mechanisms do not merely serve privacy preservation but also enhance Byzantine robustness by preventing gradient obfuscation through adaptive signal-to-noise balancing. The combination of Adaptive DP and robust aggregation enables better discrimination between benign and malicious clients.

Trade-Off Analysis of Adaptive Differential Privacy Integration

The integration of Adaptive Differential Privacy (ADP) into the DP-BREM framework introduces a critical trade-off that fundamentally shifts the mechanism from a static defense to a dynamic, performance-aware strategy, ultimately optimizing the privacy-utility frontier. The primary benefit is the substantial gain in efficiency: the Adaptive DP-BREM model achieved an overall cumulative privacy loss of ϵ=2.6 compared to $\epsilon \approx 3.9$ for the fixed-noise baseline, representing an approximate 1.3ϵ saving over 100 rounds. Moreover, this reduction in privacy cost did not compromise utility; the adaptive approach sustained superior performance, maintaining a

10–20% improvement in test accuracy over the fixed-σ model in the critical privacy budget range of ϵ=3 to ϵ=8. This gain is, however, balanced by the cost of increased algorithmic complexity and hyperparameter sensitivity on the central server, which must continuously track model performance and manage a more complex noise decay function defined by parameters such as σbase, σmin, and the scaling function κ(Acct−1). Nevertheless, the observed, measurable return on investment in both privacy budget conservation and model utility under adversarial conditions decisively justifies the marginal increase in tuning and computational overhead.

REFERENCES

1. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (pp. 308–318). https://doi.org/10.1145/2976749.2978318
2. Alistarh, D., Allen-Zhu, Z., Li, J., & Vojnović, M. (2018). Byzantine stochastic gradient descent. Advances in Neural Information Processing Systems, 31. https://proceedings.neurips.cc/paper_files/paper/2018/hash/fb9c2f44aa9d5a6fa9a595e4b089acf5-Abstract.html
3. Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS), 108, 2938–2948. https://proceedings.mlr.press/v108/bagdasaryan20a. html
4. Biggio, B., Nelson, B., & Laskov, P. (2012). Poisoning attacks against support vector machines. Proceedings of the 29th International Conference on Machine Learning (ICML-12), 1807–1814. https://proceedings.mlr.press/v28/biggio13.html
5. Blanchard, P., El Mhamdi, E. M., Guerraoui, R., & Stainer, J. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. Advances in Neural Information Processing Systems, 30. https://proceedings.neurips.cc/paper/2017/file/f4b9ec30ad9f68f89b29639786cb62ef-Paper.pdf
6. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (pp. 1175–1191). https://doi.org/10.1145/3133956.3133982
7. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konecny, J., Mazzocchi, S., McMahan, H. B., et al. (2019). Towards federated learning at scale: System design. https://arxiv.org/abs/1902.01046
8. Cao, X., Yang, M., Xie, C., & Wang, L. (2021). FLTrust: Byzantine-robust federated learning via trust bootstrapping. USENIX Security Symposium, 1429–1446. https://www.usenix.org/conference/usenixsecurity21/presentation/cao
9. Chen, L., Su, L., & Xu, J. (2017). Distributed statistical machine learning in adversarial settings: Byzantine gradient descent. ACM Measurement and Analysis of Computing Systems, 1(2), 1–25. https://doi.org/10.1145/3154503
10. Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. Theory of Cryptography Conference (pp. 265–284). https://doi.org/10.1007/11681878_14
11. Fung, C., Yoon, C. J. M., & Beschastnikh, I. (2020). The limitations of federated learning in Sybil settings. RAID 2020, 301–316. https://doi.org/10.1145/3411508.3411522
12. Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. https://arxiv.org/abs/1712.07557

13. Gu, J., Dolan-Gavitt, B., & Garg, S. (2017). BadNets: Identifying vulnerabilities in the machine learning model supply chain. https://arxiv.org/abs/1708.06733
14. Gu, Z., Wang, R., Li, J., Wang, H., & Zhu, Q. (2023). DP-BREM: Differentially private byzantine-resilient momentum aggregation for federated learning. USENIX Security Symposium 2023. https://www.usenix.org/conference/usenixsecurity23/presentation/gu-zhiqiang
15. Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2018). Federated learning for mobile keyboard prediction. https://arxiv.org/abs/1811.03604
16. Hitaj, B., Ateniese, G., & Perez-Cruz, F. (2017). Deep models under the GAN: Information leakage from collaborative deep learning. CCS '17: ACM SIGSAC Conference on Computer and Communications Security, 603–618. https://doi.org/10.1145/3133956.3134012
17. Kairouz, P., McMahan, H. B., Avent, B., et al. (2019). Advances and open problems in federated learning. https://arxiv.org/abs/1912.04977
18. Kairouz, P., McMahan, H. B., et al. (2021). Practical and responsible federated learning. Communications of the ACM, 64(12), 62–71. https://doi.org/10.1145/3454124
19. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. Proceedings of Machine Learning and Systems, 2, 429–450. https://proceedings.mlsys.org/paper/2020/file/38b9cf9b54a83ba0a512d380f6a6235f-Paper.pdf
20. Li, X., Jiang, M., Zhang, X., Kamp, M., & Dou, Q. (2021). FedBN: Federated learning on non-IID features via local batch normalization. International Conference on Learning Representations (ICLR). https://openreview.net/forum?id=6YEQUn0QICG
21. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS, 54, 1273–1282. https://proceedings.mlr.press/v54/mcmahan17a.html
22. Mohri, M., Sivek, G., & Suresh, A. T. (2019). Agnostic federated learning. International Conference on Machine Learning (ICML), 4615– 4625. https://proceedings.mlr.press/v97/mohri19a.html
23. Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. IEEE Symposium on Security and Privacy (SP), 3–18. https://doi.org/10.1109/SP.2017.41
24. Wang, Y., Yao, Y., Zheng, H., & Zhao, B. Y. (2020). A practical framework for evaluating federated learning poisoning attacks. https://arxiv.org/abs/2004.10397
25. Xie, C., Huang, K., Chen, P. Y., & Li, B. (2019). DBA: Distributed backdoor attacks against federated learning. https://arxiv.org/abs/1910.12191
26. Yin, D., Chen, Y., Kannan, R., & Bartlett, P. (2018). Byzantine-robust distributed learning: Towards optimal statistical rates. International Conference on Machine Learning (ICML), 5650–5659. https://proceedings.mlr.press/v80/yin18a.html
27. Zhang, Y., & Hu, X. (2023). Stabilizing differentially private federated learning via variance reduction. https://arxiv.org/abs/2302.01158
28. Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., & Chandra, V. (2018). Federated learning with non-IID data. https://arxiv.org/abs/1806.00582
29. Zhu, L., & Ling, H. (2022). Decoupling differential privacy and robust aggregation for better federated learning. Proceedings of the AAAI Conference on Artificial Intelligence, 36(3), 2685–2693. https://doi.org/10.1609/aaai.v36i3.20282