Privacy Protection Optimization in Federated Learning

Xinyi Zhou

College of Optoelectronics and Mechanical Engineering, Shanxi Institute of Science and Technology, Jincheng, China

Abstract. Federated learning has emerged as a promising distributed machine learning paradigm that enables collaborative model training while preserving data privacy. However, the increasing sophistication of privacy attacks and evolving regulatory requirements have exposed critical vulnerabilities in current FL systems. This paper provides a comprehensive analysis of privacy threats in federated learning, identifying three primary attack surfaces: gradient-based reconstruction, aggregation-phase breaches, and membership leakage during participant selection. This paper examines how these vulnerabilities manifest differently across healthcare, financial, and industrial applications, with sector-specific risks ranging from medical image reconstruction to inference of sensitive financial attributes. The study systematically evaluates three categories of defense mechanisms: differential privacy techniques (including adaptive noise injection and hybrid approaches), cryptographic methods (homomorphic encryption and secure multi-party computation), and blockchain-based distributed architectures. This paper analyzes the inherent trade-offs between privacy protection and model performance, presenting optimization strategies such as adaptive privacy budgeting and lightweight encryption to mitigate accuracy degradation. The paper further discusses compliance challenges posed by emerging regulations like the EU AI Act and FDA guidelines, highlighting the need for verifiable privacy proofs in sensitive domains. Finally, this paper concludes with a summary and outlook.

1 Introduction

Federated Learning (FL) represents a transformative approach to machine learning that fundamentally redefines how organizations collaborate on model development while preserving data privacy.As a distributed machine learning framework, FL enables multiple parties to jointly train models without centralizing raw data - a paradigm shift from traditional centralized approaches that require data pooling.This innovative architecture has gained substantial traction in sectors handling sensitive information, particularly healthcare and finance, where data confidentiality is both ethically imperative and legally mandated under regulations like HIPAA and GDPR.For instance, hospitals can collaboratively improve cancer detection algorithms without sharing patient records,while banks can develop better fraud detection systems without exposing customer transaction details [1].

zhouxinyi@sxist.edu.cn

The technological breakthrough of FL lies in its unique parameter exchange mechanism.Instead of transferring raw data to a central server, participants only share encrypted model updates (typically gradients or weights) during the training process [2]. This design theoretically provides inherent privacy advantages by keeping sensitive information localized on client devices.However, emerging research reveals a paradoxical reality - these very mechanisms designed to protect privacy can become vulnerabilities themselves. Sophisticated attacks can reverse-engineer sensitive data from what were presumed to be safe gradient updates. Recent studies demonstrate that determined adversaries can reconstruct high-resolution medical images from vision model gradients or infer personal attributes from model weight updates in financial applications [3].

This duality presents a critical challenge for FL adoption.On one hand, the framework offers unprecedented opportunities for privacy-preserving collaboration across institutional boundaries.Pharmaceutical companies can accelerate drug discovery by pooling knowledge without sharing proprietary compound data.National banks can collaboratively combat money laundering while maintaining customer confidentiality. On the other hand, the evolving threat landscape requires continuous advancements in privacy protection measures. Modern investigations confirm that even with rigorous safeguards like gradient clipping and noise addition, determined attackers can still extract substantial information about training data distributions through advanced statistical analysis.

This review paper systematically analyzes current research on privacy threats in federated learning, contrasting attack methodologies (membership inference, gradient inversion, and backdoor attacks) across healthcare, finance, and industrial applications. This paper critically evaluates recent studies to identify limitations in existing defense mechanisms, particularly the disconnect between theoretical privacy guarantees and practical implementation challenges. The analysis reveals inconsistent evaluation metrics across domains and highlights under-researched areas including cross-industry threat propagation and regulatory-compliance conflicts. Compared to previous surveys, this work provides the first systematic taxonomy of sector-specific vulnerabilities while synthesizing emerging protection paradigms. Our findings expose critical gaps in current literature regarding adaptive defense coordination and standardized compliance frameworks for distributed learning systems.

2 Privacy Threats in Federated Learning Systems

2.1 Comprehensive Analysis of Data Exposure Vulnerabilities

Modern federated learning (FL) systems, despite incorporating advanced security measures, continue to present multiple exploitable attack surfaces that jeopardize data privacy. In recent years have witnessed significant advancements in attack methodologies, particularly targeting gradient exchange protocols that form the backbone of FL communications. Recent studies demonstrate that malicious actors can reconstruct training samples with high precision – achieving similarity for facial recognition datasets [3].In the healthcare domain, researchers have successfully recovered chest X-rays from vision model gradients (figure 1) [4]. Similarly, in financial applications, attackers infer client income brackets with high accuracy by analyzing model weight divergences in credit scoring tasks [5].

Figure 1. From Gradient Leakage to Privacy Crisis-Exposure Risk Analysis of Chest X-ray Data in Joint Learning (SecureFed: federated learning empowered medical imaging technique to analyze lung abnormalities in chest X-rays | International Journal of Machine Learning and Cybernetics) [4]

Varous vulnerabilities manifest differently across industries. Healthcare systems are particularly susceptible to partial data reconstruction (e.g., organ segmentation in CT scans [4]), while financial institutions face attribute inference risks (e.g., loan eligibility prediction [5]). The manufacturing sector reports concerns about model fingerprinting, where industrial FL models leak equipment parameters through gradient patterns [6] .

Modern FL systems exhibit three critical vulnerability points that attackers exploit at different stages of the training pipeline. The Gradient Interface represents the most direct attack surface, where raw gradient transmission enables powerful external reconstruction attacks, particularly dangerous in healthcare applications due to its data replication potential. In parallel, the Aggregation Phase introduces distinct risks for cross-silo deployments, where man-in-the-middle attacks during model averaging can manipulate global updates without detection - a vulnerability fundamentally different from gradient leaks as it targets the collaborative process rather than individual contributions. Meanwhile, Participant Selection creates longitudinal exposure, where clients engaging in multiple rounds become susceptible to membership leakage through temporal pattern analysis, a threat that compounds over time unlike the immediate risks of gradient or aggregation attacks [3][7][8].

These vulnerabilities collectively undermine FL's core privacy guarantees through multiple attack vectors: Gradient leaks not only enable exact data replication in healthcare applications but also facilitate model inversion attacks that can compromise proprietary algorithms. Aggregation breaches permit silent model poisoning that can persist undetected across multiple training rounds, gradually degrading model performance or introducing targeted biases. Participation tracking exposes sensitive membership information through temporal analysis of update patterns, revealing which organizations or individuals contributed to specific model improvements. The compounded risks are particularly severe in cross-industry collaborations where different sectors maintain varying privacy requirements and regulatory obligations. This multifaceted threat landscape necessitates the development of sector-specific defense mechanisms that account for domain-specific data sensitivity, compliance frameworks, and adversarial capabilities while maintaining the fundamental benefits of federated learning. Current mitigation approaches must evolve beyond generic privacy protections to address these distinct attack surfaces through tailored solutions that preserve model utility without compromising security.

2.2 Evolution of Modern Attack Paradigms

Recent FL security research has identified three increasingly sophisticated threat models: membership inference attacks that exploit gradient patterns to determine training data participation, gradient inversion techniques that reconstruct raw inputs like images and text from shared updates, and distributed backdoor attacks that implant hidden triggers through poisoned local updates. These attack paradigms highlight the escalating security challenges in federated learning, where adversaries continuously develop new methods to compromise data privacy and model integrity despite the system's distributed nature.

(1) Membership Inference Attacks:

These attacks have become particularly potent in federated learning environments, where adversaries exploit gradient updates to determine if specific data samples were used in training. The attack methodology has evolved from basic statistical analysis to sophisticated techniques that leverage the unique characteristics of federated learning systems. Attackers now employ advanced correlation analysis between shared gradients and underlying data distributions, often using auxiliary datasets to train shadow models that mimic the target FL environment [9]. The temporal patterns of parameter updates in federated systems provide particularly rich leakage sources, enabling attackers to identify participation with high confidence. In healthcare applications, such techniques have been shown to successfully trace model updates back to specific hospital participants in pharmaceutical collaborations, raising serious concerns about data confidentiality in medical research partnerships.

Gradient inversion-based input reconstruction:

This technique has evolved from a theoretical threat to a practical attack, fundamentally solving an inverse optimization problem to recover high-dimensional features of raw input data from gradients. Modern approaches integrate generative adversarial network (GAN) priors, transforming gradient-matching problems into latent space searches, where candidate samples consistent with original gradients are iteratively optimized. Breakthroughs in gradient inversion techniques now all Ow near-perfect image reconstruction from minimal update rounds. Using GAN-assisted methods, attacker S can reconstruct 224×224 resolution images (comparable to ImageNet quality) within just three updat E cycles [10]. In natural language applications, approximately 60% of original text can be recovered from BERT mode L gradients, posing severe risks to confidential document processing systems [11].

(3)Distributed Backdoor Attacks:

Attackers skillfully exploit federated learning's distributed update mechanism to implant malicious behaviors by poisoning local training data or gradients to embed conditionally triggered backdoors. The novel "sleeper cell" attack pattern remains dormant until specific trigger conditions are met, making it exceptionally difficult to detect during routine evaluations. These targeted manipulations can cause the model to produce incorrect outputs exclusively for queries from specific participants.

3 Privacy-Preserving Techniques in Federated Learning

Federated learning enables collaborative model training across decentralized devices while preserving data privacy. However, privacy threats such as data leakage, inference attacks, and model poisoning necessitate robust privacy-preserving techniques. This chapter discusses three major categories of privacy protection methods: data perturbation techniques (differential privacy), cryptographic techniques (homomorphic encryption and secure multi-party computation), and distributed architecture techniques (blockchain -based privacy protection).

3.1 Data Perturbation Techniques: Differential Privacy

Differential privacy (DP) has emerged as a fundamental framework for privacy preservation in federated learning, with local differential privacy (LDP) implementations offering client-level protection. Current research demonstrates three principal technical directions in this domain. The first approach focuses on direct noise injection mechanisms for FL gradient updates. Studies have demonstrated that incorporating carefully calibrated noise mechanisms into federated learning systems can effectively balance privacy protection with model utility, as shown in comparative analyses of privacy-preserving approaches [12]. This method forms the baseline for DP in FL, though it faces challenges in high-dimensional parameter spaces where noise accumulation may significantly impact model accuracy.

Adaptive perturbation techniques represent an important advancement, dynamically adjusting noise scales based on parameter sensitivity and training stage. Research demonstrates these methods particularly improve accuracy preservation during later training phases when parameter updates stabilize [13]. The adaptive approach introduces additional computational overhead for sensitivity analysis but provides more refined privacy-utility trade-offs compared to fixed-noise methods. Implementation requires careful tuning of adaptation algorithms across different model architectures.

Building upon adaptive perturbation methods,recent hybrid approaches combine dimensionality reduction with DP mechanisms. By projecting high-dimensional gradients into lower-dimensional spaces before noise application, these methods substantially reduce required noise magnitude while maintaining equivalent privacy guarantees [14]. The technique proves particularly effective for complex models with numerous parameters, though the projection process may introduce information loss affecting convergence properties. Experimental validations show these hybrid methods can achieve comparable accuracy to non-private FL in certain scenarios while providing formal privacy proofs.However, these methods struggle with non-IID data distributions common in real-world FL scenarios, where uneven noise distribution may disproportionately degrade certain clients' model utility.

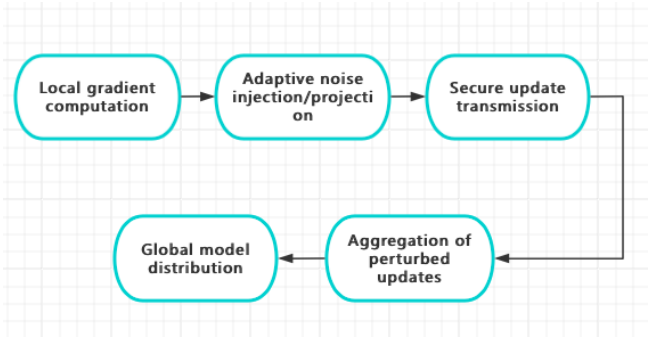

Figure 2. DP in FL Workflow (Picture credit : Original)

The evolution of DP techniques for FL shows clear progression from basic noise addition to sophisticated adaptive and hybrid methods. Current implementations demonstrate that while DP inevitably impacts model performance, careful parameterization and algorithm design can maintain usable accuracy levels for many practical applications (Figure 2). The choice between these approaches depends on specific requirements for privacy strength, computational resources, and model complexity.

3.2 Cryptographic Techniques: Homomorphic Encryption & Secure Multi-Party Computation

Cryptographic approaches in federated learning provide robust privacy protection through advanced encryption schemes that enable computation on encrypted data. Threshold-based homomorphic encryption represents a fundamental technique, where multiple participants must collaborate to decrypt aggregated results (Figure 3). This method supports secure additive operations on encrypted gradients while resisting collusion attacks, though it introduces noticeable computational overhead compared to non-encrypted operations [15].

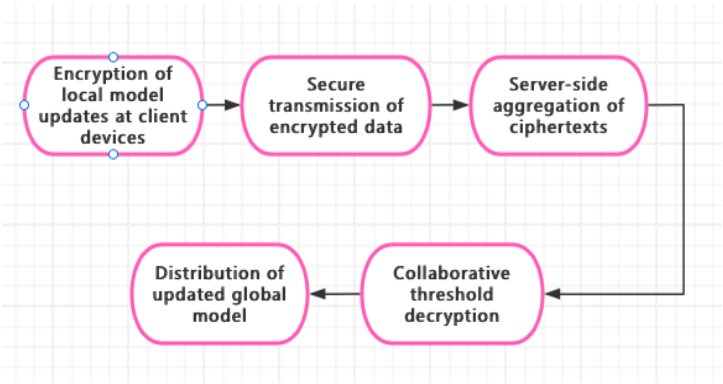

Figure 3 Cryptographic Protection Workflow (Picture credit : Original)

The development of multiparty fully homomorphic encryption has expanded these capabilities to support arbitrary computations on ciphertexts. Practical implementations employ batch processing to enhance efficiency and use approximation techniques for evaluating nonlinear functions in encrypted form. While these optimizations improve performance, they may increase communication overhead in large-scale deployments [16]. Healthcare applications have shown particular benefits from these cryptographic methods in clinical prediction scenarios. [17]

Hybrid cryptographic systems combine multiple privacy-preserving techniques to leverage their complementary strengths. These architectures often integrate homomorphic encryption with secure multi-party computation, using each method for operations where it proves most efficient. Such combinations maintain strong security guarantees while improving practical usability, though they require careful system design to coordinate different cryptographic components effectively [17]. A critical limitation surfaces in cross-silo collaborations where participants use heterogeneous encryption schemes, often requiring complex protocol conversions that introduce new attack surfaces.

3.3 Distributed Architecture Techniques: Blockchain-Based Privacy Protection

Blockchain technology enhances federated learning through decentralized architectures that provide immutable audit trails. Smart contracts manage learning processes while storing only essential verification data on-chain, effectively addressing scalability limitations [18]. The distributed ledger architecture offers additional advantages by enabling transparent verification of model updates without revealing raw data, while cryptographic hashing ensures the integrity of shared parameters (Figure 4). In healthcare applications, dual-layer architectures demonstrate significant latency reduction by separating local training from cross-institution aggregation [19]. These systems further benefit from blockchain's natural

resistance to single points of failure, which is particularly valuable in critical medical implementations where system availability is paramount.

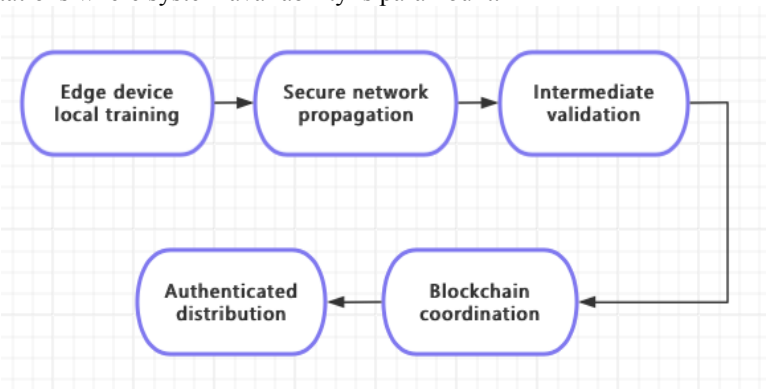

Figure 4 Distributed Architecture Workflow (Picture credit : Original)

Hierarchical access control systems implement dynamic key management with identity-based encryption and zero-knowledge proofs. These mechanisms effectively prevent unauthorized access while introducing manageable communication overhead. The integration of attribute-based access policies with blockchain smart contracts allows for fine-grained permission control that adapts to evolving security requirements across different federation participants. For enhanced security, trusted execution environments provide hardware-level protection for sensitive operations like key management [20]. This combination of cryptographic techniques and hardware security creates a defense-in-depth approach that safeguards against both external attacks and insider threats. The system's resilience is further strengthened by incorporating decentralized identity verification protocols that eliminate reliance on centralized certificate authorities. Current implementations face temporal consistency challenges in global model synchronization, particularly when handling delayed node updates in large-scale networks, which may create window periods for gradient leakage.

The choice between techniques depends on specific privacy requirements and resource constraints. Modern solutions increasingly integrate multiple approaches for comprehensive protection [19,20].

4 Privacy and Model Performance Trade-offs

In federated learning systems, the trade-off between privacy preservation and model performance constitutes a fundamental challenge. With increasingly stringent data privacy regulations and the continuous evolution of attack methods, designing solutions that simultaneously satisfy privacy requirements while maintaining model utility has become particularly critical. This trade-off stems from the inherent tension between privacy protection mechanisms (e.g., noise addition and encryption) and model learning capabilities —where stronger privacy guarantees often come at the expense of model accuracy or computational efficiency. this balance directly impacts both the practicality and regulatory compliance of federated learning systems. This chapter will systematically analyze the nature of this trade-off relationship and explore current mainstream optimization strategies, providing methodological guidance for achieving privacy-performance equilibrium across different application scenarios.

4.1 Impact of Privacy Protection on Model Performance

The implementation of privacy-preserving mechanisms in machine learning models inevitably affects their performance. In federated learning systems, the addition of differential privacy noise during gradient updates can significantly reduce model accuracy, particularly when strict privacy guarantees are required [21]. This noise injection process interferes with the model's ability to learn precise patterns from the data, leading to slower convergence rates and suboptimal final performance [22].Various privacy threats in federated learning, including membership inference and data reconstruction attacks, necessitate these protective measures despite their performance costs [23].

Privacy protection techniques like data anonymization and perturbation often degrade data utility, making it more challenging for models to extract meaningful patterns [24].The trade-off between privacy and performance becomes particularly evident in sensitive domains such as healthcare and finance, where strong privacy requirements may substantially impact model effectiveness [25]. Studies have shown that conventional privacy-preserving approaches can cause noticeable reductions in model accuracy compared to non-private alternatives[21]. However, the degree of performance degradation varies depending on the specific privacy technique employed and its implementation parameters . This creates an important research challenge: how to maintain adequate privacy protection while minimizing negative impacts on model utility [23].

The relationship between privacy and performance isn't linear - small increases in privacy protection can sometimes lead to disproportionately large performance drops . This nonlinearity suggests that careful optimization of privacy parameters is crucial for achieving an acceptable balance [25]. Recent work has demonstrated that the performance impact of privacy mechanisms can be mitigated through intelligent algorithm design and parameter tuning [21]. Understanding these privacy-performance dynamics is essential for developing practical privacy-preserving machine learning systems .

4.2 Optimization Strategies

4.2.1 Adaptive Privacy Budget

Adaptive privacy budget allocation has emerged as an effective strategy for balancing privacy and performance. Unlike traditional approaches that distribute privacy budgets uniformly across iterations, adaptive methods dynamically adjust the budget based on gradient magnitudes [21]. During early training stages when gradients are typically larger, these methods allocate smaller privacy budgets (permitting more noise), while later stages receive larger budgets for finer parameter tuning . This approach recognizes that not all training iterations equally impact final model performance .

The DP-AGD algorithm demonstrates how adaptive budgeting can improve model accuracy compared to fixed allocation methods [24]. By concentrating privacy resources where they're most needed, such algorithms achieve better final performance while maintaining the same overall privacy guarantees. Multi-tier privacy frameworks extend this concept by applying different protection levels based on data sensitivity . For instance, highly sensitive personal data might receive stronger protection than less sensitive metadata [22].

Implementation of adaptive privacy budgets requires careful consideration of several factors. The rate of budget adjustment must balance immediate privacy needs with long-term model convergence [23]. Some approaches use gradient magnitude thresholds to trigger budget reallocation, while others employ more sophisticated statistical measures [24]. These methods have shown particular promise in federated learning scenarios, where

data heterogeneity makes fixed privacy budgets particularly problematic [25]. The effectiveness of adaptive budgeting also depends on the specific differential privacy variant being implemented, with some variants being more amenable to dynamic allocation than others [21].

4.2.2 Lightweight Encryption

Lightweight encryption techniques offer a practical compromise between computational overhead and privacy protection. These methods focus on securing the most vulnerable aspects of model training while minimizing performance impacts [21]. Gradient sparsification, for example, reduces encryption costs by only protecting the most significant gradient components [22]. This selective protection maintains model accuracy while significantly decreasing computation and communication costs compared to full encryption [23].

Partial homomorphic encryption represents another lightweight approach that enables specific operations on encrypted data [24]. By supporting only necessary mathematical operations (like addition), these schemes achieve practical performance while still providing meaningful privacy guarantees [25]. In federated learning systems, such techniques can protect gradient updates without requiring the full complexity of fully homomorphic encryption . The choice of encryption parameters (like key size and sparsification level) allows tuning of the privacy-performance trade-off [22].

Blockchain-based privacy solutions provide an alternative lightweight approach through decentralized verification [23]. Rather than encrypting all data, these systems use cryptographic hashes and consensus mechanisms to ensure data integrity and prevent unauthorized modifications [24]. While not providing the same level of privacy as full encryption, blockchain approaches can be sufficient for many applications while offering better scalability . The effectiveness of lightweight encryption depends heavily on proper implementation - overly aggressive optimization can leave vulnerabilities, while excessive caution negates the performance benefits .

4.2.3 Joint Optimization

Joint optimization approaches integrate privacy preservation directly into the model training process rather than treating it as a separate component [21]. These methods simultaneously optimize for both model performance and privacy protection, often through carefully designed objective functions . One common strategy combines differential privacy with gradient sparsification, reducing the amount of noise needed while maintaining privacy guarantees [23]. This joint approach has demonstrated improved accuracy compared to sequential privacy application [24].

The CARISMA tool exemplifies how joint optimization can reduce implementation costs by identifying privacy issues early in the design process [25]. By analyzing system models for potential privacy violations during development, this approach prevents costly post-hoc modifications [21].Other joint methods incorporate privacy constraints directly into the learning algorithm's update rules [22]. These constrained optimization techniques can achieve better privacy-performance trade-offs than simple noise addition [23].

Federated learning systems particularly benefit from joint optimization approaches . Techniques that co-optimize client selection, local computation, and privacy preservation can significantly improve overall system efficiency [25]. Some advanced methods even adapt their optimization strategy based on real-time assessments of privacy risk and model performance . The effectiveness of joint optimization depends on properly balancing the various objectives - overemphasizing privacy can hurt performance, while neglecting

privacy can lead to vulnerabilities [22]. Recent work has shown that carefully designed joint optimization frameworks can approach non-private model performance while still providing strong privacy guarantees [23].

5 Conclusion

Federated learning has emerged as a promising paradigm for collaborative model training while preserving data privacy across decentralized devices and institutions. However, as this paper has demonstrated, modern FL systems face significant privacy threats that undermine their core security guarantees. In response to these challenges, this paper has examined three major categories of privacy-preserving techniques. Data perturbation methods, particularly differential privacy, have evolved from basic noise injection to sophisticated adaptive and hybrid approaches that better balance privacy and utility. Cryptographic techniques, including homomorphic encryption and secure multi-party computation, provide robust protection through advanced encryption schemes, though with notable computational overhead. Distributed architecture techniques leveraging blockchain technology offer decentralized solutions with immutable audit trails and enhanced resistance to single points of failure. The implementation of these privacy mechanisms inevitably impacts model performance, creating fundamental trade-offs that require careful optimization. Our analysis reveals that privacy protection measures can reduce model accuracy, slow convergence rates, and increase computational costs. However, strategic approaches like adaptive privacy budgeting, lightweight encryption, and joint optimization methods can mitigate these impacts while maintaining adequate protection. These optimization strategies demonstrate that intelligent algorithm design and parameter tuning can achieve acceptable balances between privacy and performance.

Looking forward, federated learning systems must continue evolving their privacy protections to keep pace with advancing attack methodologies. Future research should focus on developing more efficient implementations of existing techniques, exploring novel hybrid approaches that combine multiple protection layers, and creating standardized frameworks for privacy certification and audit. The ultimate goal remains achieving robust privacy preservation without unduly compromising model utility or system performance.

Reference

1. Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., & Cardoso, M. J.: 'The future of digital health with federated learning', NPJ Digital Medicine, 2020, 3, pp. 119.
2. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A.: 'Communication-efficient learning of deep networks from decentralized data', In Artificial intelligence and statistics (pp. 1273-1282). PMLR, April 2017.
3. Geiping, J., Bauermeister, H., Dröge, H., & Moeller, M.: 'Inverting gradients-how easy is it to break privacy in federated learning'. Advances in neural information processing systems, 2020, 33, pp. 16937-16947.
4. Makkar, A., & Santosh, K. C.: 'SecureFed: federated learning empowered medical imaging technique to analyze lung abnormalities in chest X-rays', International Journal of Machine Learning and Cybernetics, 2023, 14(8), pp. 2659-2670.
5. Exploiting unintended feature leakage in collaborative learning. In 2019 IEEE symposium on security and privacy (SP) (pp. 691-706). IEEE.

6. Wang, H., Eklund, D., Oprea, A., & Raza, S.: 'FL4IoT: IoT device fingerprinting and identification using federated learning', ACM Transactions on Internet of Things, 2023, 4(3), pp. 1-24.
7. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S.,& Seth, K.: 'Practical secure aggregation for federated learning on user-held data'. arXiv preprint arXiv:1611.04482, 2016.
8. Nasr, M., Shokri, R., & Houmansadr, A.: 'Comprehensive privacy analysis of deep learning'. Proceedings of the 2019 IEEE Symposium on Security and Privacy (SP), Vol. 2018, pp. 1-15.
9. Little, C., Elliot, M., & Allmendinger, R.: 'Federated learning for generating synthetic data: a scoping review', International Journal of Population Data Science, 2023, 8(1), pp. 2158.
10. Yin, H., Mallya, A., Vahdat, A., Alvarez, J. M., Kautz, J., & Molchanov, P.: 'See through gradients: Image batch recovery via gradinversion'. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 16337-16346), 2021.
11. Feng, T., Hashemi, H., Hebbar, R., Annavaram, M., & Narayanan, S. S.: 'Attribute inference attack of speech emotion recognition in federated learning settings'. arXiv preprint arXiv:2112.13416, 2021.
12. Zheng, H., Hu, H., & Han, Z.: 'Preserving user privacy for machine learning: Local differential privacy or federated machine learning?', IEEE Intelligent Systems, 2020, 35(4), pp. 5-14.
13. Truex, S., Liu, L., Chow, K. H., Gursoy, M. E., & Wei, W.: 'LDP-Fed: Federated learning with local differential privacy'. In Proceedings of the third ACM international workshop on edge systems, analytics and networking (pp. 61-66), April 2020.
14. Wang, Y. X., Lei, J., & Fienberg, S. E.: 'Learning with differential privacy: Stability, learnability and the sufficiency and necessity of ERM principle', Journal of Machine Learning Research, 2016, 17(183), pp. 1-40.
15. Ronald Cramer, Ivan Damgärd, Jesper Buus Nielsen. "Multiparty Computation from Threshold Homomorphic Encryption"
16. Mouchet, C. V.: 'Multiparty homomorphic encryption: From theory to practice', Doctoral dissertation, EPFL, 2023.
17. Kumar, A. V., Sujith, M. S., Sai, K. T., Rajesh, G., & Yashwanth, D. J. S.: 'Secure Multiparty computation enabled E-Healthcare system with Homomorphic encryption'. In IOP Conference Series: Materials Science and Engineering, Vol. 981, No. 2, p. 022079, December 2020.
18. Feng, Q., He, D., Zeadally, S., Khan, M. K., & Kumar, N.: 'A survey on privacy protection in blockchain system', Journal of network and computer applications, 2019, 126, pp. 45-58.
19. Qi, M., Wang, Z., Han, Q. L., Zhang, J., Chen, S., & Xiang, Y.: 'Privacy protection for blockchain-based healthcare IoT systems: A survey', IEEE/CAA Journal of Automatica Sinica, 2022.
20. Ma, M., Shi, G., & Li, F.: 'Privacy-oriented blockchain-based distributed key management architecture for hierarchical access control in the IoT scenario', IEEE Access, 2019, 7, pp. 34045-34059.

21. Binjubeir, M., Ahmed, A. A., Ismail, M. A. B., Sadiq, A. S., & Khan, M. K.: 'Comprehensive survey on big data privacy protection', IEEE Access, 2019, 8, pp. 20067-20079.
22. Lee, J., & Kifer, D.: 'Concentrated differentially private gradient descent with adaptive per-iteration privacy budget', In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1656-1665), July 2018.
23. Ahmadian, A. S., Strüber, D., Riediger, V., & Jürjens, J.: 'Supporting privacy impact assessment by model-based privacy analysis'. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing (pp. 1467-1474), April 2018.
24. Novado, D., Cohen, E., & Foster, J.: 'Multi-tier privacy protection for large language models using differential privacy'. Authorea Preprints, 2024.
25. Solove, D. J., & Hoofnagle, C. J.: 'A model regime of privacy protection', U. Ill. L. Rev., 2006, pp. 357.