Privacy-Utility Trade-offs in Federated Learning for 6G Networks: A Systematic Evaluation of Software-based Privacy Mechanisms

Jawaad Ahmar*[1], Iqra Batool*[2], Mostafa M. Fouda† ‡[3], Mohamed I. Ibrahem§[4], and Zubair Md Fadlullah*[5]
*Department of Computer Science, Western University, London, ON, Canada.
†Department of Electrical and Computer Engineering, Idaho State University, Pocatello, ID, USA.
‡Center for Advanced Energy Studies (CAES), Idaho Falls, ID, USA.
§School of Computer and Cyber Sciences, Augusta University, Augusta, GA, USA.
Emails: [1]jahmar@uwo.ca [2]ibatool2@uwo.ca, [3]mfouda@ieee.org, [4]mibrahem@augusta.edu, [5]zfadlullah@ieee.org

***Abstract*—Federated Learning (FL) enhances privacy by training models locally, but remains vulnerable to inference attacks. We systematically evaluate software-based privacy mechanisms to address the question: "How does applying Homomorphic Encryption (HE), Differential Privacy (DP), and hybrid approaches in Federated Learning affect model accuracy, privacy protection, and resource overhead in 6G networks?" We implemented five configurations: Baseline FL, FL+HE, FL+DP, Standard Hybrid (simultaneous HE+DP), and our novel Sequential Hybrid (DP during training, HE during aggregation). Evaluations on CIFAR-10 using our reproducible Docker-based framework revealed clear trade-offs: accuracy declined progressively from baseline (68.34%) to HE (57.90%), DP (35.05%), and hybrid approaches (33-35%), while privacy improved correspondingly. Our sequential hybrid approach improved learning stability compared to simultaneous application. HE demonstrated superior resilience to data heterogeneity, maintaining 92.1% accuracy under non-IID conditions versus 56-66% for hybrid approaches. This systematic evaluation provides empirical guidance for selecting privacy mechanisms based on specific 6G network requirements, where both strong privacy and performance across heterogeneous devices are essential.**

***Index Terms*—6G Networks ,Federated learning, differential privacy, homomorphic encryption, privacy-utility trade-off, data heterogeneity**

I. Introduction

Machine learning models increasingly process sensitive personal data, raising critical privacy concerns in emerging 6G networks, which will support 1000× higher data rates and 10× lower latency than 5G, enabling ubiquitous AI across trillions of devices [1]. Federated Learning (FL) enhances privacy by training models locally, but model updates remain vulnerable to inference attacks [2], potentially compromising future 6G applications. Current privacy approaches include hardware-based Trusted Execution Environments (TEEs) and software-based mechanisms. Hardware solutions face significant limitations—many devices lack necessary hardware, secure memory is limited, and side-channel vulnerabilities exist [3]. These constraints are particularly problematic for heterogeneous 6G networks [4], necessitating scalable software-based solutions like Homomorphic Encryption (HE) and Differential Privacy (DP). HE enables computation on encrypted data [5], ideal for FL but computationally expensive, while DP adds calibrated noise [6], providing formal privacy guarantees but reducing model accuracy. Existing research typically evaluates these mechanisms in isolation, leaving potential synergies unexplored, especially for resource-constrained edge devices in 6G networks [7]. This paper systematically investigates software-based privacy mechanisms in FL, addressing: "How does applying HE, DP, and combined approaches in FL affect model accuracy, privacy protection, and resource overhead in high-performance networks?" We address three critical research gaps: (1) lack of consistent experimental frameworks for direct comparison, (2) unexplored potential of software-based hybrid approaches for resource-constrained devices, and (3) limited evaluation of privacy mechanisms under non-IID data distributions—crucial for heterogeneous 6G networks. Figure 1 illustrates the fundamental components of federated learning, where heterogeneous client devices train local models on their private data and communicate model updates via high-speed 6G infrastructure to a central server for aggregation.

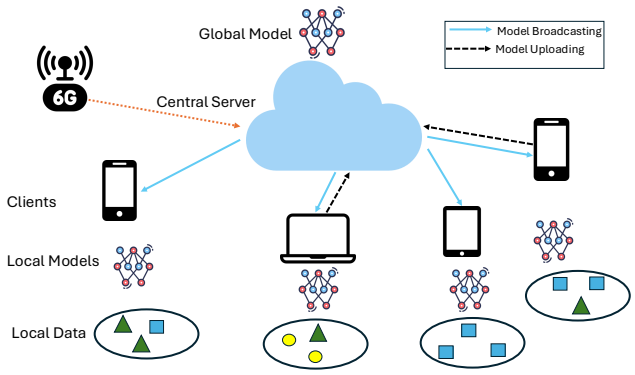

Fig. 1. Federated Learning Architecture with Heterogeneous Data Distribution

We implement and evaluate five configurations (baseline

FL, FL+HE, FL+DP, standard hybrid, and our novel sequential hybrid approach) using consistent metrics across both IID and non-IID data distributions. Our findings reveal trade-offs between privacy and utility, with accuracy declining from baseline (68.34%) to HE (57.90%), DP (35.05%), and hybrid approaches (33-35%). Importantly, HE maintains 92.1% accuracy under non-IID conditions while hybrid approaches retain only 56-66% [8]. Our most significant contribution is a novel sequential hybrid approach applying DP during training and HE during aggregation. This temporal separation improves learning stability and heterogeneity resilience, addressing 6G requirements for strong privacy and high model utility in heterogeneous environments [9].

The remainder of this paper is organized as follows. Section II provides background on federated learning, privacy vulnerabilities, and existing privacy mechanisms, along with related work in the context of 6G networks. Section III details our methodology, including system architecture, implementation of privacy mechanisms, and experimental design. Section IV presents our experimental results, analyzing the trade-offs between privacy, utility, and resource efficiency across different configurations. Section V discusses the implications of our findings, limitations, and suggests future research directions. Finally, Section VI concludes the paper with a summary of contributions and their significance for privacy-preserving federated learning in future 6G networks.

II. Related Work

A. Federated Learning (FL)

Federated Learning enables collaborative model training without centralizing sensitive data [1]. Clients train on private data and share only model updates with a server through the Federated Averaging (FedAvg) algorithm [10]. While this approach minimizes data movement, model updates can still leak information about training data [11].

B. Privacy Vulnerabilities in FL

Despite structural privacy advantages, FL remains vulnerable to sophisticated attacks. Membership Inference Attacks (MIAs) determine whether specific records were used during training. Nasr et al. [2] demonstrated that both passive and active white-box inference attacks can succeed against federated models, with privacy leakage occurring in communications and the final model.

C. Privacy-Enhancing Techniques in FL

Homomorphic Encryption (HE) enables computation on encrypted data [3], allowing encrypted model updates, but introduces substantial computational overhead. Differential Privacy (DP) provides formal guarantees by adding calibrated noise to computations [6]. Abadi et al. introduced DP-SGD [5], and Geyer et al. extended this to federated learning with client-level DP [8]. The key challenge is balancing privacy with model utility. Hardware-based solutions using Trusted Execution Environments face practical limitations for heterogeneous deployments [9], [12].

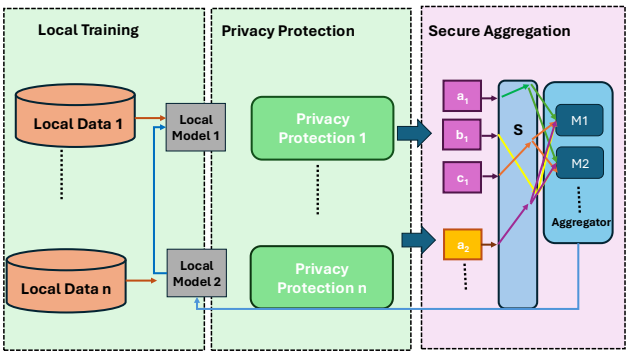

Fig. 2. Privacy Preserving FL Framework

D. FL Privacy in 6G Networks

6G networks introduce new privacy considerations. Yang et al. [13] and Liu et al. [14] highlight privacy as critical for FL integration within 6G architecture. Letaief et al. [15] emphasize the need for adaptive privacy approaches across distributed intelligence layers. Saad et al. [16] note the necessity for customizable privacy guarantees across diverse stakeholders. Porambage et al. [17] highlight the tension between 6G's sub-millisecond latency requirements and computationally intensive privacy mechanisms. Current research gaps include: (1) lack of systematic comparison within consistent experimental frameworks, (2) insufficient exploration of software-based combined approaches, (3) limited understanding of performance under non-IID data distributions, and (4) absence of research considering 6G's specific requirements. Our research addresses these gaps by evaluating five privacy approaches within a unified framework.

III. Methodology

We implemented a comprehensive federated learning framework to systematically evaluate software-based privacy mechanisms across utility, privacy, and efficiency dimensions, with particular consideration for 6G network constraints.Figure 2 illustrates the Privacy Preserving Federated learning(FL) framework.

A. Federated Learning System Architecture

Our system follows a client-server architecture using PyTorch, with two primary components. The server component orchestrates the FL process by distributing the global model, aggregating client updates using FedAvg, and evaluating model performance. It also implements server-side privacy operations, including encrypted aggregation for HE-based approaches. The client component is responsible for local model training, applying client-side privacy mechanisms, and secure communication with the server. Each client operates independently on their local data partition.

For neural network architecture, we implemented two CNN variants optimized for different privacy mechanisms.

SimpleCNN, a three-layer convolutional network followed by two fully connected layers, was used for baseline and DP experiments. SmallCNN, a streamlined two-layer network with reduced parameters, was optimized for computationally intensive HE operations. Both networks target CIFAR-10 image classification using cross-entropy loss and SGD optimization (momentum=0.9, learning rate=0.01).

B. Privacy Mechanisms Implementation

1) Differential Privacy (DP): We integrated Opacus (a production-ready DP library for PyTorch) implementing the following components:

- **Gradient Clipping**: Limits sensitivity by constraining per-sample gradients to a maximum L2 norm of 1.0
- **Noise Addition**: Applies calibrated Gaussian noise using configurable noise multipliers (0.3-1.0)
- **Privacy Accounting**: Accurately tracks privacy budget (ε) using Rényi Differential Privacy accounting

2) Homomorphic Encryption (HE): We leveraged Pyfhel (a Python interface for Microsoft SEAL) implementing the CKKS scheme with:

- **Parameter Quantization**: Converts floating-point model parameters to fixed-point representation with configurable bit precision (16-24 bits)
- **Chunking**: Segments large parameter tensors to accommodate CKKS polynomial degree limitations
- **Encrypted Aggregation**: Performs secure addition on encrypted model updates without decryption

Our configuration used polynomial modulus degree 4096, coefficient modulus bits [40, 20, 40], and scale 2^{20}, balancing security, precision, and computational efficiency.

3) Hybrid Approaches: We implemented two novel hybrid combinations of DP and HE:

- **Standard Hybrid**: Simultaneously applies both mechanisms, with DP during training and HE for encrypting the already-noised model updates
- **Sequential Hybrid**: Temporally separates the mechanisms by first applying DP during local training, then separately applying HE during aggregation with optimized parameters (noise multiplier 0.3, quantize bits 24)

C. Experimental Design

We employed a factorial experimental design to systematically evaluate privacy mechanisms across multiple dimensions:

- **Privacy Configurations**: Baseline FL (no privacy), FL+HE, FL+DP, Standard Hybrid, Sequential Hybrid
- **Data Distributions**: IID and non-IID (Dirichlet allocation with $\alpha = 0.5$ and $\alpha = 0.1$)
- **System Parameters**: 10 clients, 10-20 training rounds, batch size 64, local epoch 1

For 6G relevance, we focused on heterogeneous data distributions and resource constraints typical of edge computing environments. All experiments were containerized using Docker to ensure reproducibility and facilitate deployment.

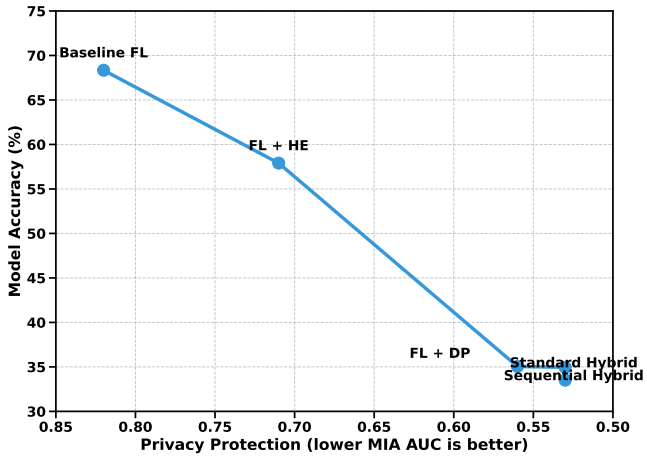

Fig. 3. Privacy-Utility Trade-off

D. Evaluation Methodology

We evaluated each configuration using multi-faceted metrics aligned with critical 6G requirements:

- **Utility Metrics**: Model accuracy and loss on a separate test set, with particular attention to convergence rate
- **Privacy Metrics**: Vulnerability to membership inference attacks (MIA), measured by attack AUC, precision, recall, and accuracy
- **Resource Metrics**: Computation time per round, communication overhead, and memory usage
- **Convergence Behavior**: Learning stability and round-to-round improvement patterns

For MIA evaluation, we implemented both threshold-based and advanced shadow model attacks following established methodologies. We conducted multiple trials with different random seeds to ensure statistical validity of results.

This comprehensive evaluation framework enables direct comparison of privacy mechanisms under consistent conditions, addressing a critical research gap in the current literature.

IV. Results

Our systematic evaluation revealed fundamental trade-offs between privacy, utility, and resilience in federated learning systems. This section presents our empirical findings across multiple dimensions.

A. Privacy-Utility Trade-off Analysis

Our results demonstrated a clear inverse relationship between privacy protection and model utility, as shown in Table I. Figure 3 visualizes this inverse relationship between privacy protection and model utility, clearly demonstrating how stronger privacy guarantees come at the cost of decreased accuracy

TABLE I
FINAL TEST ACCURACY AND PRIVACY PROTECTION COMPARISON (IID SETTING)

Approach	Final Accuracy (%)	Privacy Protection (MIA AUC)
Baseline FL	68.34	0.82
FL + HE	57.90	0.71
FL + DP	35.05	0.56
FL + Standard Hybrid	34.95	0.53
FL + Sequential Hybrid	33.51	0.53

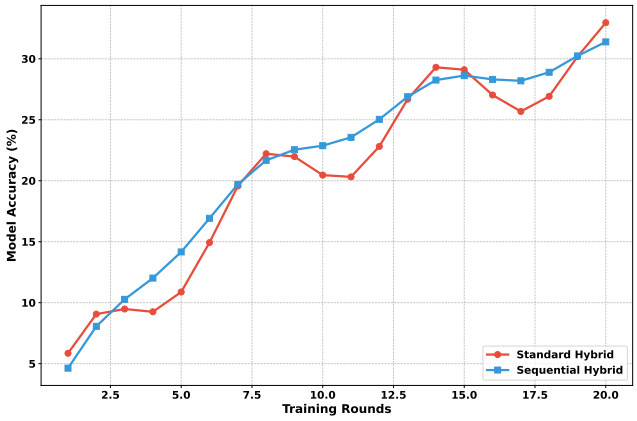

Fig. 4. Learning Progression Analysis

The baseline approach achieved the highest accuracy (68.34%) but offered minimal privacy protection (MIA AUC of 0.82). As we applied increasingly robust privacy mechanisms, accuracy progressively decreased while privacy improved. Homomorphic encryption reduced accuracy by approximately 10 percentage points (to 57.90%) while providing moderate privacy improvement. In contrast, differential privacy caused a substantially larger accuracy reduction (to 35.05%) but offered significantly stronger privacy guarantees. The hybrid approaches achieved similar final accuracy to DP alone but provided slightly better privacy protection, with MIA AUC values of 0.53.

B. Learning Progression Analysis

The learning progression across training rounds revealed important behavioral differences between privacy mechanisms, illustrated in Figure 4. Figure 3 compares the learning trajectories of standard and sequential hybrid approaches across training rounds, demonstrating the more consistent improvement and reduced fluctuations achieved through temporal separation of privacy mechanisms Baseline and HE approaches demonstrated faster convergence with stable learning trajectories. In contrast, DP-based methods exhibited slower initial progress and more erratic convergence patterns, likely due to the added noise interfering with gradient optimization.

Our sequential hybrid approach demonstrated more consistent improvement with fewer fluctuations compared to the standard hybrid approach. This supports our hypothesis that temporal separation of privacy mechanisms prevents error compounding and leads to more stable learning.

C. Resource Overhead Analysis

Table II summarizes the computational and communication requirements for each privacy approach.

HE-based approaches introduced significant computational overhead, increasing the per-round processing time by approximately 3× compared to baseline. DP added relatively modest computational overhead (about 30% increase). In terms of communication efficiency, HE-based approaches used 94% less bandwidth compared to baseline and DP approaches. However, this efficiency stemmed from the necessary use of smaller model architectures for computationally intensive approaches rather than an inherent property of the encryption method itself.

D. Sequential vs. Standard Hybrid Approach

Our novel sequential hybrid approach demonstrated several advantages over the standard approach. It achieved higher starting accuracy (14.38% vs. 10.12% after round 1), more consistent round-to-round improvements, smaller variance in validation accuracy, and better resilience to parameter changes. The improved performance can be explained by examining how privacy mechanisms introduce error: DP adds calibrated noise to gradients, potentially pushing optimization away from the true minimum, while HE introduces approximation errors through parameter quantization. In the standard approach, these errors compound—HE approximation errors affect already-noised gradients. The sequential approach isolates these error sources, preventing error multiplication and resulting in more stable learning. Figure 4 demonstrates this stability difference, showing how our sequential approach maintains more consistent improvement with fewer fluctuations compared to the standard hybrid method.

E. Impact of Data Heterogeneity

Table III summarizes how each approach performed under different data distributions. Figure 5 illustrates the differential resilience of privacy mechanisms to data heterogeneity, highlighting how HE maintains over 92% of its accuracy under highly non-IID conditions, while hybrid approaches exhibit substantially greater sensitivity.

We discovered that privacy mechanisms exhibit differential resilience to data heterogeneity. HE demonstrated the

TABLE II
RESOURCE REQUIREMENTS COMPARISON

Approach	Time per Round (s)	Communication (MB)	Total Time (s)
Baseline FL	9.45	1750.89	94.5
FL + HE	28.76	104.43	287.6
FL + DP	12.38	1750.89	123.8
FL + Standard Hybrid	32.18	104.43	321.8
FL + Sequential Hybrid	29.54	104.43	295.4

TABLE III
IMPACT OF DATA HETEROGENEITY ON ACCURACY

Approach	IID Accuracy (%)	Non-IID (α=0.5)	Non-IID (α=0.1)	Retention (%)
Baseline FL	68.34	66.06	58.17	85.1
FL + HE	57.90	54.65	53.32	92.1
FL + DP	35.05	32.28	28.51	81.3
FL + Standard Hybrid	34.95	23.95	19.52	55.9
FL + Sequential Hybrid	33.51	30.78	22.25	66.4

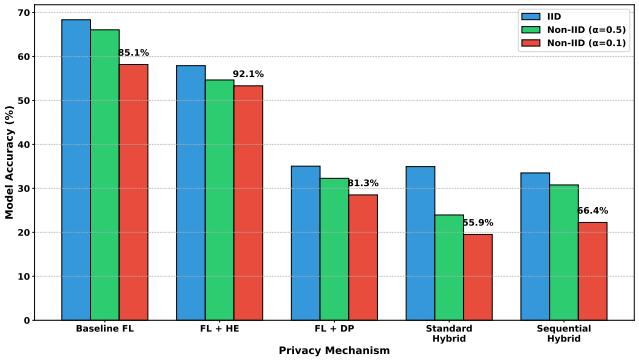

Fig. 5. Impact of Data Heterogeneity on Accuracy

strongest resilience, retaining 92.1% of its IID accuracy under high heterogeneity (α=0.1), while baseline and DP showed moderate impact, retaining 85.1% and 81.3% of their accuracy, respectively. Hybrid approaches were most vulnerable to heterogeneity, with the standard hybrid retaining only 55.9% of its accuracy. The sequential hybrid improved resilience over the standard hybrid approach, retaining 66.4% of its accuracy. This pattern suggests that approaches with more complex privacy mechanisms become increasingly sensitive to data heterogeneity, likely because non-IID data introduces optimization challenges that are exacerbated by privacy-preserving perturbations.

F. Three-Way Trade-off Analysis

1) Privacy-Utility-Communication Trade-off: Our multi-dimensional analysis revealed that homomorphic encryption provides unique benefits in the three-dimensional space of privacy, utility, and communication efficiency:

- **Baseline**: High utility, low privacy, high communication cost
- **HE**: Good utility, moderate privacy, excellent communication efficiency
- **DP**: Lower utility, good privacy, high communication cost
- **Hybrid approaches**: Lower utility, best privacy, excellent communication efficiency

2) Privacy-Utility-Heterogeneity Trade-off: We observed a significant inverse relationship between privacy strength and heterogeneity resilience. Approaches with stronger privacy guarantees demonstrated markedly greater sensitivity to data distribution shifts. The standard hybrid approach exhibited approximately 31% accuracy loss when moving from IID to non-IID data, nearly six times the heterogeneity sensitivity of HE.

This suggests a fundamental tension between privacy and heterogeneity resilience that has not been widely recognized in privacy-preserving federated learning research. The sequential hybrid approach offers a valuable compromise, providing strong privacy protection while demonstrating better heterogeneity resilience than the standard hybrid approach.

V. DISCUSSION

Our systematic evaluation of privacy mechanisms in federated learning revealed critical insights about trade-offs between privacy, utility, and resilience. Here we examine these findings and their implications for future research.

A. Validity Considerations

Our results should be interpreted in light of certain limitations. Internally, our specific parameter selections for privacy mechanisms may influence outcomes, though sensitivity analyses mitigated this concern. We leveraged established libraries (Opacus, Pyfhel) and conducted multiple trials with different random seeds to reduce stochastic effects. Externally, our focus on CIFAR-10 and simplified CNN architectures within a simulated federated environment may not fully represent all domains or real-world 6G deployments. Our privacy evaluation primarily relied on MIA success metrics, which provide one perspective on privacy protection.

B. Key Insights and Implications

Our research reveals that sequential application of privacy mechanisms (DP during training, HE during aggregation) outperforms simultaneous implementation, challenging conventional approaches to privacy-preserving FL. We identified a critical inverse relationship between privacy strength and heterogeneity resilience—HE maintains 92.1% accuracy under non-IID conditions versus only 56-66% for hybrid approaches, a crucial finding for heterogeneous 6G environments. For practical applications, mechanism selection should be requirement-driven: HE for high-utility scenarios with moderate privacy needs; DP or hybrid approaches when privacy is paramount; and our sequential hybrid approach when balancing privacy with heterogeneity resilience.

6G networks will likely require tiered privacy approaches based on device capabilities and application contexts, with our sequential hybrid method specifically addressing the tension between privacy requirements and heterogeneous data distribution. Future 6G architectures should incorporate these privacy-utility-heterogeneity trade-offs as core design parameters, supporting adaptive mechanisms that dynamically adjust to data distribution changes, heterogeneity-aware privacy solutions, and energy-efficient implementations suitable for resource-constrained edge devices.

VI. Conclusion

Our research systematically evaluated software-based privacy mechanisms in federated learning, revealing fundamental trade-offs between privacy, utility, and data heterogeneity resilience. We quantified how privacy mechanisms progressively reduce model accuracy while improving privacy protection, with HE causing moderate accuracy impact (10.4 percentage points), DP causing larger reductions (33.3 points), and hybrid approaches providing the strongest privacy at the highest accuracy cost. The most significant contribution is our novel sequential hybrid approach that applies DP during training and HE during aggregation. This temporal separation improved learning stability and heterogeneity resilience compared to standard simultaneous application, challenging conventional wisdom in privacy-preserving machine learning. We discovered that privacy mechanisms vary significantly in their resilience to data heterogeneity, with HE maintaining 92.1% of its accuracy under highly non-IID conditions while hybrid approaches retain only 56-66Future research should focus on developing adaptive privacy mechanisms that dynamically adjust parameters based on data distribution and convergence, creating heterogeneity-aware privacy solutions specifically designed for non-IID distributions, establishing theoretical frameworks for sequential privacy applications, and exploring energy-efficient implementations suitable for resource-constrained edge devices in 6G networks. These directions will help realize the vision of privacy-aware distributed intelligence in future ultra-high-performance networks where both strong privacy guarantees and high model utility across heterogeneous environments are essential.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[2] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in *2019 IEEE symposium on security and privacy (SP)*. IEEE, 2019, pp. 739–753.

[3] Z. Gu, H. Huang, J. Zhang, D. Su, H. Jamjoom, A. Lamba, D. Pendarakis, and I. Molloy, "Confidential inference via ternary model partitioning," *arXiv preprint arXiv:1807.00969*, 2018.

[4] S. Moriai, "Privacy-preserving deep learning via additively homomorphic encryption," in *2019 IEEE 26th Symposium on Computer Arithmetic (ARITH)*. IEEE Computer Society, 2019, pp. 198–198.

[5] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016, pp. 308–318.

[6] M. Naehrig, K. Lauter, and V. Vaikuntanathan, "Can homomorphic encryption be practical?" in *Proceedings of the 3rd ACM workshop on Cloud computing security workshop*, 2011, pp. 113–124.

[7] I. Batool, M. M. Fouda, and Z. M. Fadlullah, "Deep learning-based throughput prediction in 5g cellular networks," in *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*. IEEE, 2024, pp. 1–6.

[8] R. C. Geyer, T. Klein, and M. Nabi, "Differentially private federated learning: A client level perspective," *arXiv preprint arXiv:1712.07557*, 2017.

[9] B. Jayaraman and D. Evans, "Evaluating differentially private machine learning in practice," in *28th USENIX Security Symposium (USENIX Security 19)*, 2019, pp. 1895–1912.

[10] W. N. Khan, H. Siddiqui, A. K. Das, V. Patel, M. Shafi, and J. H. P. Yoon, "Federated learning for 6g-enabled secure communication systems: a comprehensive survey," *Artificial Intelligence Review*, vol. 56, pp. 15 525–15 608, 2023.

[11] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Exploiting unintended feature leakage in collaborative learning," in *2019 IEEE symposium on security and privacy (SP)*. IEEE, 2019, pp. 691–706.

[12] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, and Y. Liu, "{BatchCrypt}: Efficient homomorphic encryption for {Cross-Silo} federated learning," in *2020 USENIX annual technical conference (USENIX ATC 20)*, 2020, pp. 493–506.

[13] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial intelligence-enabled intelligent 6g networks," *IEEE Network*, vol. 34, no. 6, pp. 272–280, 2020.

[14] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6g communications: Challenges, methods, and future directions," *China Communications*, vol. 17, no. 9, pp. 105–118, 2020.

[15] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6g: Ai empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, 2019.

[16] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[17] P. Porambage, G. Gür, D. P. M. Osorio, M. Liyanage, A. Gurtov, and M. Ylianttila, "The roadmap to 6g security and privacy," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1094–1122, 2021.