Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 19 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i19.57619

PAPER

Advanced Privacy-Utility Optimization Techniques in Federated Learning with Differential Privacy for IoMT – A Review

Shaista Ashraf Farooqi[1] (✉), Aedah Abd Rahman[1], Amna Saad[2]

[1]Asia e University (AeU), Subang Jaya, Selangor, Malaysia

[2]Universiti Kuala Lumpur, Malaysian Institute of Information Technology, Kuala Lumpur, Malaysia

shaista@safarooqi.com

ABSTRACT

This paper reviews advanced optimization techniques to address the privacy-utility tradeoff in federated learning with differential privacy (FL-DP), focusing on applications in the Internet of Medical Things (IoMT). IoMT systems face significant challenges, including heterogeneous, non-IID data distributions, resource-constrained devices, and stringent privacy regulations such as HIPAA and GDPR, making it complex to ensure robust privacy while maintaining high model utility. The review explores methods such as adaptive privacy budgeting, which dynamically adjusts privacy parameters ($\in$) based on data sensitivity and device capabilities, and client selection strategies that enhance global model accuracy by prioritizing high-quality data contributions while effectively managing privacy budgets. Techniques like gradient clipping and noise scaling are examined for their ability to mitigate the negative impact of differential privacy (DP) noise, ensuring stability in real-time applications like remote patient monitoring and anomaly detection. This study analyzes existing techniques and identifies gaps in advancing scalable and efficient FL-DP frameworks in IoMT. Future directions include AI-driven adaptive privacy mechanisms and energy-efficient optimization algorithms to enhance the scalability, performance, and sustainability of FL-DP in IoMT environments. These advancements aim to develop secure, high-performance IoMT systems that comply with privacy standards while addressing real-world healthcare challenges.

KEYWORDS

federated learning with differential privacy (FL-DP), security and privacy, Internet of Medical Things (IoMT), optimization techniques, client selection, privacy-utility tradeoff

1 INTRODUCTION

The widespread adoption of the Internet of Medical Things (IoMT) has revolutionized healthcare by enabling seamless, decentralized, and real-time data acquisition from various interconnected medical devices. These devices, from

Farooqi, S. A., Rahman, A. A., Saad, A. (2025). Advanced Privacy-Utility Optimization Techniques in Federated Learning with Differential Privacy for IoMT – A Review. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(19), pp. 134–150. https://doi.org/10.3991/ijim.v19i19.57619

Article submitted 2025-07-09. Revision uploaded 2025-08-19. Final acceptance 2025-08-19.

wearables to remote diagnostic tools, generate substantial volumes of sensitive and personal health data, raising significant concerns about privacy and security [1]. Ensuring the privacy of such sensitive information is paramount, especially given the stringent regulatory frameworks, such as HIPAA and GDPR, which mandate robust privacy measures. This paper examines advanced optimization techniques within the federated learning with differential privacy (FL-DP) framework, tailored for the IoMT. While FL-DP provides a robust solution for ensuring data privacy in distributed learning, achieving an optimal privacy-utility tradeoff remains a significant challenge, particularly in resource-constrained IoMT environments characterized by non-IID data distributions. Figure 1 illustrates various applications of the Internet of Medical Things.

Fig. 1. IoMT applications

1.1 Differential privacy-based federated learning framework

Federated learning (FL) has emerged as a crucial technique to facilitate privacy-preserving distributed machine learning by allowing data to remain localized on devices while sharing only model updates. Integrating differential privacy (DP) has been proposed to further strengthen privacy in FL, adding a layer of protection by ensuring that individual data points cannot be reconstructed or inferred from model updates. The FL-DP framework introduces noise into the model updates, which, while enhancing privacy, presents significant challenges in maintaining model utility—a critical factor in IoMT applications, where accuracy and timely insights can directly affect healthcare outcomes [2] [3]. The privacy-utility tradeoff in FL-DP remains a fundamental issue, as increasing privacy through noise addition often leads to a reduction in model accuracy, which can be detrimental in real-time applications such as remote patient monitoring or anomaly detection.

1.2 Privacy-utility tradeoff

Privacy-utility concerns and data heterogeneity are key challenges in IoMT systems, where data collected from various medical devices can vary widely in terms of distribution and quality. Non-IID data distributions can hinder the performance of global models in FL. The privacy-utility tradeoff in IoMT is critical for balancing the need for data privacy with the accuracy of machine learning models used in healthcare applications. As IoMT devices collect sensitive medical data, privacy regulations such as HIPAA and GDPR require robust protections, such as DP [4] [5]. However, increasing privacy by adding noise to data can reduce model utility, leading to less accurate predictions in real-time health applications.

1.3 Non-IID data distribution

The non-IID nature of IoMT data arises from variations in device types, user behaviors, and environmental conditions, posing significant challenges to achieving a compelling privacy-utility tradeoff in FL-DP frameworks. This data heterogeneity often disrupts model training, leading to imbalanced contributions and degraded global model performance. Furthermore, resource-constrained IoMT devices are burdened by the additional computational and communication requirements of privacy-preserving mechanisms, further complicating the tradeoff. Ensuring an optimal balance between robust data privacy and high model utility is critical for delivering accurate and reliable healthcare outcomes [6]. Achieving this balance enables IoMT systems to comply with privacy regulations such as HIPAA and GDPR while operating efficiently within the constraints of the devices.

1.4 Resource constraints

Implementing FL-DP in the IoMT is challenging due to devices' limited computational power, memory, and battery life. These constraints are further strained in large-scale IoMT networks, where frequent model updates and privacy-preserving mechanisms demand significant resources. Consequently, developing scalable and efficient FL-DP frameworks for IoMT devices is complex. Achieving an optimal privacy-utility tradeoff under such conditions becomes increasingly complex, as ensuring robust privacy often increases computational and energy costs. Overcoming these challenges requires developing advanced optimization techniques that address energy efficiency, minimize communication overhead, and enhance scalability [62]. Such solutions must balance privacy protection with model performance while enabling IoMT devices to operate effectively within resource constraints, ensuring reliable and secure deployments in real-time healthcare applications [7].

The study highlights adaptive privacy budgeting and client selection strategies to optimize privacy levels and improve global model performance. Techniques like gradient clipping and noise scaling are evaluated to reduce the impact of DP noise on accuracy, while strategies for minimizing communication overhead and enhancing energy efficiency are emphasized for scalable IoMT deployments. The structure of the paper is as follows: Section 2 presents the related work, Section 3 provides a detailed discussion, Section 4 outlines future directions, and Section 5 concludes the paper, followed by the references.

2 RELATED WORK

Federated learning with DP enables collaborative machine learning across distributed medical devices while preserving the confidentiality of sensitive health data. This approach allows IoMT devices to train shared models without exchanging raw data, ensuring patient privacy and regulation compliance. By integrating FL and DP, healthcare systems can leverage diverse datasets to improve diagnostic accuracy and patient outcomes without compromising data security. FL allows IoMT devices, such as wearable sensors and diagnostic tools, to train models locally without transferring raw data to a central server, thus minimizing the risk of exposing personal health information [8]. Incorporating DP enhances this framework by introducing mathematically rigorous noise into model updates, ensuring that individual data points remain indistinguishable, even when aggregated in global models [9].

However, FL DP presents unique technical challenges in IoMT, particularly in balancing the privacy-utility tradeoff. The noise introduced to protect privacy can adversely affect model performance, which is especially critical in real-time IoMT applications, such as continuous patient monitoring and anomaly detection [10]. Furthermore, the heterogeneous nature of IoMT data—often non-IID across devices—combined with the limited computational resources of IoMT devices adds to the complexity of ensuring efficient learning while adhering to privacy standards. Therefore, advanced optimization techniques are required to address these challenges, ensuring robust privacy guarantees without compromising model accuracy, scalability, or energy efficiency in resource-constrained IoMT environments [11] [12].

2.1 FL-DP implementation phases

The implementation of the FL-DP framework involves a systematic process to ensure data privacy while maintaining model utility. Each phase is designed to address specific challenges in IoMT systems, such as data heterogeneity, resource constraints, and compliance with privacy regulations [13] [14]. The following seven steps outline the key stages in implementing FL-DP, ensuring a robust, privacy-preserving, and scalable framework for IoMT applications.

Phase 1: Data collection and initialization. Internet of Medical Things devices collect decentralized, sensitive data, which remains stored locally to ensure privacy. The central server initializes a global model and shares it with all participating devices. This phase establishes the foundation for collaborative learning while adhering to privacy principles.

Phase 2: Local model training. Each device trains the global model locally using its private dataset. DP mechanisms, such as gradient clipping and noise addition, are applied during training to protect individual data contributions. This ensures data privacy even during computation.

Phase 3: Local update protection. Before sending updates to the central server, devices apply DP techniques to secure their contributions. Noise is added to the model gradients or updates, and clipping is used to limit the sensitivity of the updates, ensuring compliance with privacy budgets.

Phase 4: Secure aggregation. The central server aggregates the DP-protected updates using cryptographic methods such as secure multiparty computation (SMPC) or Homomorphic Encryption. These techniques ensure that individual contributions remain confidential during the aggregation process.

Phase 5: Global model update. The aggregated updates are used to refine the global model. This phase involves monitoring the consumption of the privacy budget to maintain DP guarantees. The updated global model incorporates learnings from all participating devices.

Phase 6: Model distribution. The refined global model is shared back with the devices for the next training round. This iterative process continues until the global model achieves the desired level of performance or convergence.

Phase 7: Deployment and evaluation. The final global model is validated on test data to ensure accuracy and robustness. Once validated, it is deployed in real-world IoMT applications such as patient monitoring, anomaly detection, or diagnostic support. This phase marks the practical implementation of the FL-DP framework in healthcare environments.

These steps ensure that the FL-DP framework is implemented efficiently, balancing privacy guarantees, resource constraints, and model utility in IoMT systems. Figure 2 shows the seven phases of the FL-DP framework.

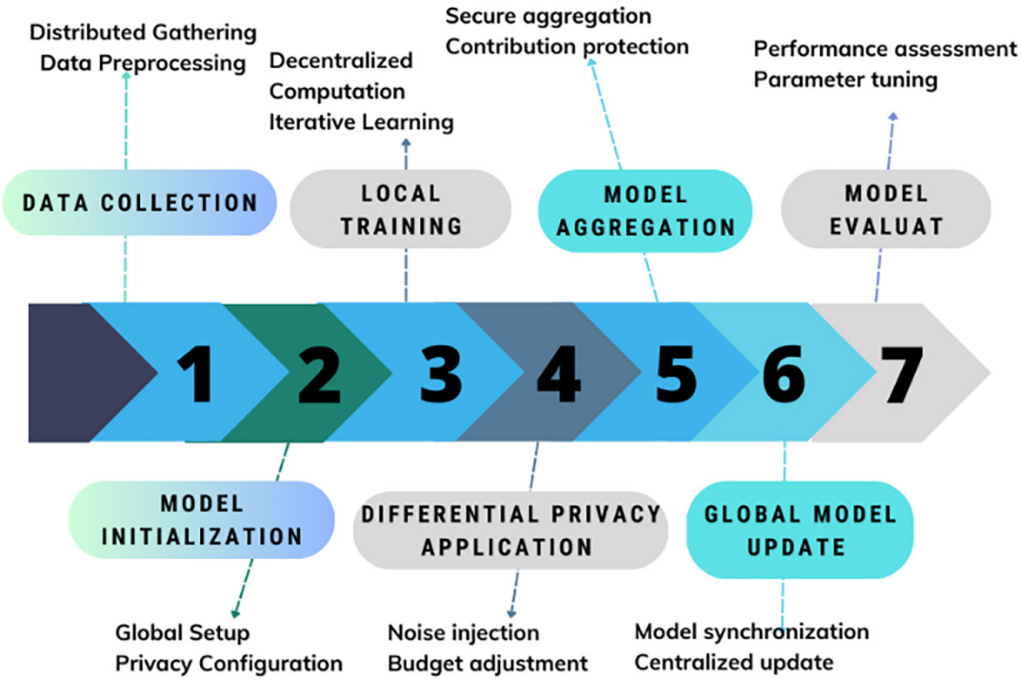

Fig. 2. 7-phases of the FL-DP framework

2.2 Optimization techniques

In the FL-DP framework, numerous advanced optimization techniques have been engineered to mitigate the intrinsic privacy-utility tradeoff and enhance overall model efficiency. These techniques are specifically designed to optimize the performance of FL models while safeguarding the confidentiality of distributed data, particularly in resource-constrained environments such as the IoMT. Given the unique challenges presented by IoMT, including limited computational resources, non-IID data distributions, and stringent privacy requirements, these optimization strategies have become crucial in ensuring both privacy preservation and model accuracy [15] [16]. The following outlines key high-impact optimization techniques

that have been integrated into FL DP frameworks to address these challenges and improve scalability, utility, and security in such decentralized systems.

Adaptive privacy budgeting. Adaptive privacy budgeting revolutionizes DP in FL, particularly in dynamic environments like the IoMT. It fine-tunes privacy budgets during training by adjusting noise levels based on model status and data sensitivity. Unlike static methods, this approach adapts in real-time, applying more noise during high-risk phases and reducing it as the model stabilizes, thus improving accuracy while maintaining privacy [17]. This method is vital for scalable FL in healthcare, minimizing overhead without sacrificing privacy. Integrating feedback loops further refines privacy-utility tradeoffs, enhancing performance in critical tasks like predictive health analytics and anomaly detection in IoMT [18].

Gradient clipping and noise scaling. Gradient clipping is a crucial mechanism in FL, imposing upper bounds on the magnitude of model gradients from individual clients. It constrains the potential influence of any single client on the global model during aggregation, mitigating the risk of outliers or adversarial updates disproportionately affecting the model. When gradient clipping is integrated with noise scaling, the amount of noise added to each gradient is dynamically adjusted relative to the clipped gradient's magnitude. It ensures that the noise injected to uphold DP scales appropriately as the model progresses towards convergence [19]. In the context of noise scaling, the noise level is correspondingly reduced as the model approaches an optimal state, where gradients naturally decrease in magnitude. This proportional noise reduction maintains stringent privacy protections and minimizes model performance degradation, leading to a more optimal balance between privacy guarantees and model utility [20]. As a result, the privacy-utility tradeoff is refined, especially in resource-constrained environments like the IoMT, where maintaining high utility without compromising privacy is critical to the system's effectiveness.

Client selection. Client selection in FL is pivotal in optimizing the overall model performance by selectively involving clients that provide the most valuable data contributions. Since not all clients contribute equally due to the heterogeneity of data distribution across devices, advanced techniques such as contribution-based or adaptive client selection are employed. These methods prioritize clients based on data quality, diversity, and computational resources, ensuring that those with high-quality or more representative datasets are chosen to participate in each training round [21]. These selection strategies reduce the inclusion of clients with noisy or low-quality data, minimizing unnecessary DP noise. This strategic client participation enhances model convergence, improving the overall privacy-utility tradeoff.

Secure aggregation. Secure aggregation, a critical technique in FL, relies on the use of advanced cryptographic methods such as homomorphic encryption and multi-party computation (MPC). These methods are instrumental in ensuring that individual client updates remain encrypted while enabling their secure aggregation at the server. Homomorphic encryption allows mathematical operations on encrypted data without decryption, thereby protecting the model parameters. Similarly, MPC enables multiple parties to compute a function while keeping their inputs private collaboratively, further enhancing the confidentiality of each client's data during aggregation [22]. Secure aggregation reduces the need for additional DP noise, as the privacy of individual client contributions is inherently protected. As a result, the overall noise addition required to meet privacy guarantees is minimized, thereby preserving the model's utility while ensuring robust privacy protection [23]. These cryptographic approaches are particularly valuable in sensitive applications like the IoMT, where maintaining data confidentiality and model accuracy is paramount.

Hybrid differential privacy. Hybrid differential privacy (Hybrid DP) combines both centralized (CDP) and local differential privacy (LDP) mechanisms to optimize the privacy-utility tradeoff across heterogeneous client environments. By applying LDP directly at the client level, sensitive data is protected before transmission, while CDP ensures broader privacy protection at the server during aggregation. This dual-layer approach allows for tailored privacy budget allocation, adjusting privacy mechanisms based on client data sensitivity and resource availability, which is crucial in resource-constrained environments like IoMT. Hybrid DP provides enhanced scalability, efficiency, and adaptive privacy protections within distributed learning systems [24] [25].

Communication-efficient learning. Communication-efficient FL leverages advanced techniques such as compressed communication and threshold-based updates to optimize network and computational resources in IoMT environments. Compressed communication reduces the data payload by encoding model gradients or updates using techniques like quantization or sparsification, minimizing the transmission size without significantly degrading information content. Threshold-based updates further enhance efficiency by ensuring that only model updates surpassing a predefined magnitude threshold are transmitted, reducing the frequency of communication and conserving bandwidth [26] [27] [28] [60]. These methods improve communication efficiency, energy consumption, and model accuracy, making them vital in resource-constrained, latency-sensitive FL deployments.

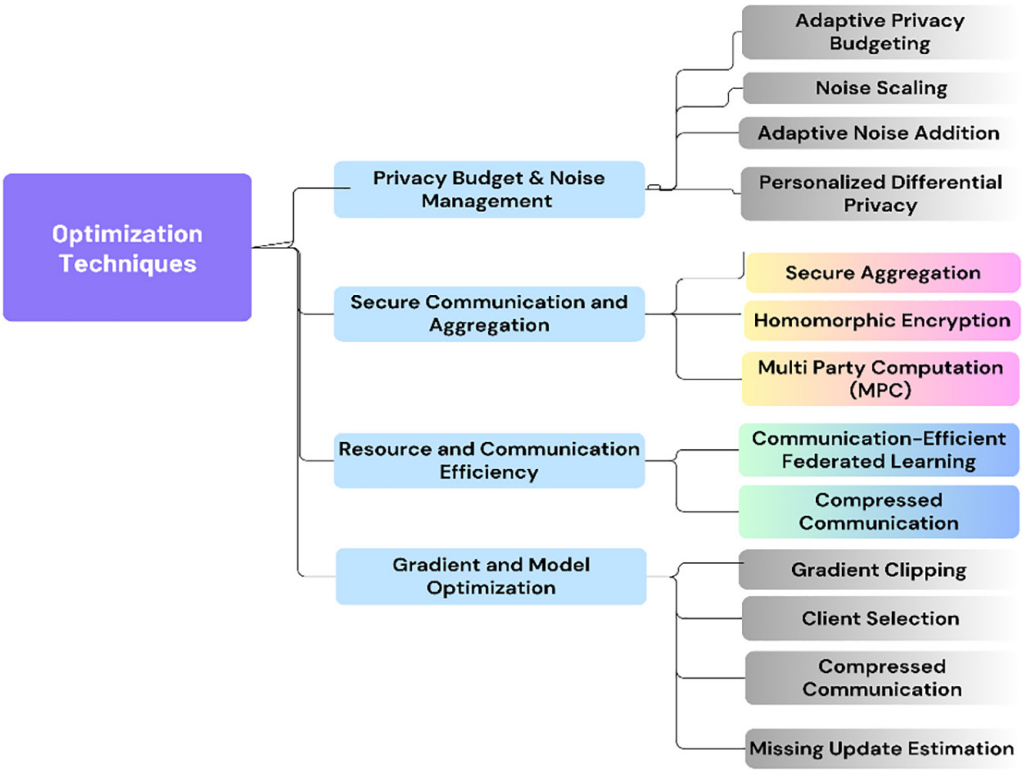

Fig. 3. Optimization techniques

The related work highlights the foundational phases of FL-DP implementation and various optimization techniques designed to address the critical privacy-utility tradeoff. These studies provide essential insights into balancing privacy preservation and model performance, offering a solid foundation for advancing

scalable, efficient, and privacy-compliant frameworks tailored to the unique challenges of IoMT systems. Figure 3 illustrates the key areas of optimization techniques.

3 DISCUSSION

The discussion examines the challenges of balancing privacy and utility in the IoMT systems. It highlights diverse data distributions, limited device resources, and communication overhead. The section critiques current FL-DP methods, noting their effects on model performance, scalability, and efficiency. By identifying these challenges, the discussion suggests areas for improvement and potential solutions to enhance IoMT system performance and privacy compliance.

3.1 Challenges in enhancing privacy-utility balance in IoMT systems

The IoMT produces vast amounts of sensitive biomedical data, requiring advanced privacy-preserving techniques to maintain confidentiality while ensuring data utility. Achieving an optimal privacy-utility balance is challenging due to non-IID data, device heterogeneity, and resource constraints like limited bandwidth and energy. DP and FL offer promising solutions for secure, distributed data processing in IoMT systems. However, scaling these methods in dynamic, large-scale IoMT networks introduces trade-offs in computational overhead, communication latency, and cryptographic complexity. Addressing these challenges is essential for IoMT systems to provide real-time healthcare services while complying with strict privacy standards.

Intrinsic trade-offs between privacy and utility. Balancing privacy and utility in IoMT systems is challenging, as DP protects data by adding noise, which can reduce model accuracy. In IoMT applications such as clinical diagnostics, where precision is critical, managing the privacy budget ($\in$) becomes crucial. A lower $\in$ provides stronger privacy guarantees but results in a decline in model utility, while a higher $\in$ improves utility at the expense of weaker privacy protections [29].

In FL-based IoMT systems, the trade-off becomes more complex as noise is added to local updates from heterogeneous, non-IID data sources. At the same time, devices face resource and communication limitations. Advanced approaches, such as Rényi DP and adaptive privacy budget allocation, have been developed to mitigate the loss of utility while ensuring privacy [30]. Furthermore, emerging techniques like gradient clipping, privacy amplification, and personalized DP offer promising solutions to address these challenges, facilitating secure and high-performance IoMT systems in healthcare settings.

Heterogeneity of data and non-IID distributions. Internet of Medical Things systems generate non-IID data from diverse sources like wearable sensors, implantable devices, and EHRs, with significant variability across devices. Such heterogeneity introduces significant challenges for applying uniform privacy-preserving mechanisms, particularly those based on DP. Uniform noise injection in diverse datasets can imbalance the privacy-utility trade-off, degrading model accuracy or leaving sensitive data vulnerable to attacks.

To effectively address this complexity, developing adaptive privacy-preserving techniques is critical. These methods use dynamic privacy budget allocation, adjusting the privacy parameter ($\in$) based on data sensitivity, distribution, or model needs [31]. Such approaches enable fine-grained control over noise applications, optimizing the balance between privacy and utility based on the statistical properties of the data.

FL-based IoMT systems need client-specific DP strategies to handle non-IID data and varying device capabilities. Furthermore, advanced techniques such as adaptive gradient clipping, personalized DP, and per-client noise calibration can ensure that the varying sensitivity and heterogeneity of IoMT data are accommodated without compromising model performance or privacy guarantees [32].

Diverse in device resources and client participation. Internet of Medical Things systems include devices with varying capabilities; high-performance devices can handle complex algorithms, while resource-constrained ones may struggle due to limited power, memory, and bandwidth. This heterogeneity complicates privacy-preserving mechanisms, as resource-limited devices cannot implement advanced techniques like differential privacy, causing imbalances in privacy and utility. In FL-based IoMT architectures, where model training is distributed across client devices, not all devices may participate in every training round due to varying availability or resource limitations. This partial client participation, often referred to as client drop-out or client selection variability, directly influences the convergence rate and performance of the global model [33] [34].

To address this, advanced techniques such as personalized federated learning (PFL) and hierarchical federated learning (HFL) enable differential aggregation tailored to each device's capabilities. Reinforcement learning-based client selection and adaptive noise scaling help improve model updates and privacy for diverse clients. Dynamic privacy budgeting, gradient clipping, and SMPC enable low-power IoMT devices to participate securely and efficiently. These methods ensure robust privacy and performance in complex, high-dimensional IoMT networks [35].

Communication overhead and system efficiency. Privacy-preserving techniques, particularly in FL frameworks, often require increased communication between devices and the central server to ensure that privacy constraints are met. Additional communication increases energy use and bandwidth in resource-limited IoMT devices, affecting performance and delaying real-time healthcare processing. Top of Form

The communication overhead affects the system's performance and poses a challenge for maintaining utility, as frequent communication delays can impede real-time processing, which is often essential in healthcare applications. Consequently, optimizing communication protocols to ensure privacy without compromising system efficiency and utility remains a pressing concern [36] [37].

Adversarial attacks on privacy-preserving models. Despite applying privacy-preserving techniques like differential privacy, IoMT systems remain vulnerable to adversarial attacks, such as model inversion and reconstruction attacks. In these scenarios, an adversary attempts to reconstruct sensitive data (e.g., patient health records) from the model's outputs or intermediate updates. Protecting against such attacks requires enhancing privacy guarantees, which can further reduce the model's utility. Additionally, FL-based IoMT systems are susceptible to inference attacks, where an adversary infers sensitive information from partial model updates sent by individual clients. Designing robust security mechanisms against these sophisticated attacks while ensuring that model performance and utility are not unduly compromised is a significant challenge in IoMT privacy research [38] [39] [40] [61].

Scalability in large-scale IoMT networks. Scaling privacy-preserving mechanisms in large IoMT networks is challenging due to device capabilities, activity levels, and non-IID data differences. Techniques like DP and FL require significant computational and communication resources. Secure multiparty computation (SMPC) and homomophobic encryption (HE) add complexity, leading to delays in low-power devices and bandwidth-constrained networks, impacting real-time healthcare applications.

Advanced methods are key to ensuring high performance and strong data security in large IoMT networks. Techniques like model compression, pruning, and quantization reduce communication overhead, while hierarchical FL with edge computing improves aggregation efficiency. Dynamic privacy budget allocation enhances real-time network performance. Emerging solutions like blockchain-based decentralized FL further support scalable, privacy-preserving IoMT systems [41] [42] [43]. Table 1 presents the challenges in enhancing the privacy-utility balance within IoMT systems, alongside corresponding potential solutions.

Table 1. Challenges of privacy utility tradeoff

Challenges	Description	Potential Solution
Intrinsic Trade-off Between Privacy and Utility	Balancing data privacy with the utility of IoMT systems is challenging, as enhancing privacy often reduces data utility.	– Implement Differential Privacy (DP) to add controlled noise, protecting individual data while maintaining overall data utility. – Use Federated Learning (FL) to train models across decentralized devices without sharing raw data, preserving privacy and utility.
Heterogeneity of Data and Non-IID Distributions	IoMT devices generate diverse, non-independent, and identically distributed (non-IID) data, complicating unified data analysis.	– Develop personalized FL models that account for individual device data characteristics. – Apply clustering techniques to group similar data distributions, enabling more effective model training.
Scalability in Large-Scale IoMT Networks	Managing privacy and utility across extensive IoMT networks with numerous devices poses significant challenges.	– Develop hierarchical FL architectures to manage large-scale device networks effectively. – Utilize edge computing to process data locally, reducing the burden on central servers and enhancing scalability.
Diverse Device Resources and Client Selection	IoMT devices vary in computational power, memory, and energy capacity, affecting participation in FL processes.	– Implement adaptive client selection algorithms that consider device capabilities and data quality. – Utilize lightweight cryptographic protocols to accommodate resource-constrained devices.
Communication Overhead	Frequent data exchanges in FL can lead to high communication costs, especially for bandwidth-limited IoMT devices.	– Employ communication-efficient protocols, such as gradient compression and quantization, to reduce data transmission volumes. – Schedule periodic aggregation to minimize communication frequency

3.2 Discussion on privacy-utility optimization algorithms

Several advanced algorithms have been proposed to optimize the privacy-utility tradeoff in FL-DP, particularly in the context of IoMT systems. These algorithms aim to address the unique challenges posed by resource-constrained devices, non-IID data distributions, and the need for real-time healthcare applications.

Adaptive privacy budget allocation. One widely used approach is adaptive privacy budgeting, which dynamically adjusts the privacy budget ($\in$) based on the data's sensitivity and the model's requirements. This algorithm ensures a balanced tradeoff between privacy and utility by allocating higher privacy budgets to less sensitive data or critical updates. While effective, its implementation can increase computational complexity, particularly in large-scale IoMT networks with heterogeneous devices [44] [45].

Gradient clipping. Gradient clipping limits the magnitude of gradients during training to manage the sensitivity of updates. This technique simplifies the addition of DP noise and ensures stable model convergence, especially in environments with non-IID data. However, excessive clipping can distort gradient updates, impacting model accuracy, particularly in scenarios requiring high precision, such as medical diagnostics [46].

Noise scaling algorithms. Noise scaling dynamically adjusts the amount of noise added to updates based on the number of participating clients or the sensitivity of gradients. This method leverages privacy amplification by sub-sampling, reducing the noise burden when more clients are involved. Despite its advantages, noise scaling can introduce variability in privacy guarantees, particularly in IoMT networks with fluctuating client participation [47] [48].

Reinforcement learning-based client selection. To enhance model utility, reinforcement learning-based client selection identifies clients based on data quality, computational resources, and reliability. This strategy minimizes the inclusion of low-quality updates and improves global model performance. However, its reliance on learning policies can increase system overhead and complicate scalability in dynamic IoMT environments [49] [50].

Federated model pruning and quantization. Model pruning and quantization optimize communication and computation by reducing the size of model updates. These techniques are particularly beneficial in IoMT systems with bandwidth and energy constraints. While effective in reducing overhead, they may lead to slight accuracy drops, necessitating careful parameter tuning to maintain performance [51] [52].

Personalized Differential Privacy (PDP). Personalized DP customizes privacy budgets based on individual client data sensitivity and contributions. This approach maximizes utility while ensuring strong privacy for sensitive datasets. However, implementing PDP in IoMT systems with large-scale, heterogeneous clients can introduce fairness concerns and computational burdens [53] [54].

Table 2. Various optimization algorithms

Optimization Algorithm Category	Algorithm	Key Approaches	Challenges	IoMT Use Cases
Adaptive Privacy Budget Allocation	DPAdaMod_AGC [55]	Utilizes adaptive gradient clipping to improve accuracy without compromising privacy.	Determining optimal clipping thresholds and privacy budgets.	Protecting sensitive patient data during model training.
Gradient Clipping	DP-SGD [56]	Incorporates gradient clipping to control sensitivity before adding noise, ensuring differential privacy in stochastic gradient descent.	Balancing noise addition with model accuracy.	Ensuring privacy in distributed medical data analysis.
Noise Scaling Algorithms	Adap DP-FL [57]	Uses multi-agent reinforcement learning for dynamic privacy budget allocation.	Computational complexity and convergence issues	Optimizing client participation in health monitoring networks.
Federated Model Pruning and Quantization	QuanCrypt-FL [58]	Combines structured and unstructured pruning for personalized federated learning under data heterogeneity.	Maintaining model performance post-pruning.	Enhancing efficiency in wearable medical device networks.
Personalized Differential Privacy	Pldp-fl [59]	Adjusts privacy mechanisms based on the context and sensitivity of data, offering personalized privacy guarantees.	Implementation complexity and ensuring consistent privacy guarantees.	Tailoring privacy levels in personalized healthcare applications.

These algorithms provide diverse strategies to optimize the privacy-utility tradeoff in FL-DP systems for IoMT. However, their effectiveness is often limited by computational complexity, scalability challenges, and system heterogeneity. Further research is needed to develop lightweight, adaptive, and scalable solutions that address the dynamic requirements of IoMT systems while ensuring robust privacy and utility.

4 FUTURE DIRECTION

Advancing FL-DP in IoMT systems necessitates innovative optimization techniques to balance data privacy and utility. Future research should focus on adaptive algorithms, efficient communication methods, and personalized models to enhance performance and scalability in these complex environments.

4.1 Adaptive optimization algorithms

Developing algorithms that dynamically adjust privacy parameters, such as the privacy budget ($\in$), is essential for enhancing the privacy-utility trade-off. By tailoring privacy settings based on data sensitivity, model convergence, and device constraints, these algorithms minimize the negative impact of noise on model performance, ensuring efficiency and scalability.

4.2 Gradient optimization techniques

Improving gradient optimization methods, such as gradient clipping and noise reduction strategies, is crucial for reducing accuracy degradation while preserving privacy. Advanced gradient descent techniques, including momentum-based and adaptive optimizers, can accelerate convergence and improve model performance in FL-DP frameworks.

4.3 Client selection optimization

Client selection will play a pivotal role in enhancing FL-DP performance by prioritizing clients with high-quality data, significant contributions, and sufficient resources. This approach improves global model accuracy, computational efficiency, and privacy guarantees, ensuring robust and reliable IoMT applications.

4.4 Communication efficiency

Optimizing communication costs between IoMT devices and servers is critical for large-scale deployments. Techniques like gradient compression, quantization, and sparsification can significantly reduce bandwidth usage and latency while maintaining model performance and privacy.

4.5 Personalized optimization techniques

Personalized optimization strategies focus on tailoring models to individual IoMT devices while preserving differential privacy. These methods address data heterogeneity and varying device capabilities, ensuring models are better suited to specific use cases and improving overall system utility.

These future directions provide a road map for advancing optimization techniques in FL-DP frameworks, addressing key challenges in IoMT systems. Together, these strategies will enhance the performance, scalability, and privacy compliance of IoMT applications, paving the way for efficient and secure healthcare systems.

5 CONCLUSION

In conclusion, this review has examined the critical challenges in optimizing the privacy-utility trade-off within FL-DP for IoMT systems. We identified key obstacles such as data heterogeneity, resource constraints, and communication overhead by analyzing existing literature. Our review highlights advanced techniques proposed to address these issues, including adaptive privacy mechanisms, communication-efficient protocols, and personalized DP models. We also discussed optimization methods like gradient adjustment, client selection, and noise reduction, which are integral to enhancing model accuracy while maintaining stringent privacy guarantees. As IoMT systems continue to evolve, the insights from this review provide valuable guidance for researchers and practitioners aiming to develop secure, efficient, and high-utility healthcare applications. Ongoing research into these optimization strategies will advance the field and ensure robust privacy protections in future IoMT deployments.

6 REFERENCES

[1] F. Kamalov, B. Pourghebleh, M. Gheisari, Y. Liu, and S. Moussa, "Internet of Medical Things privacy and security: Challenges, solutions, and future trends from a new perspective," *Sustainability*, vol. 15, no. 4, p. 3317, 2023. https://doi.org/10.3390/su15043317

[2] R. Hu, Y. Guo, H. Li, Q. Pei, and Y. Gong, "Personalized federated learning with differential privacy," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9530–9539, 2020. https://doi.org/10.1109/JIOT.2020.2991416

[3] Y. Zhang, Y. Lv, and F. Liu, "A systematic survey for differential privacy techniques in federated learning," *Journal of Information Security*, vol. 14, no. 2, pp. 111–135, 2023. https://doi.org/10.4236/jis.2023.142008

[4] J. Zhou, Z. Su, J. Ni, Y. Wang, Y. Pan, and R. Xing, "Personalized privacy-preserving federated learning: Optimized trade-off between utility and privacy," in *GLOBECOM 2022–2022 IEEE Global Communications Conference*, 2022, pp. 4872–4877. https://doi.org/10.1109/GLOBECOM48099.2022.10000793

[5] M. Kim, O. Günlü, and R. F. Schaefer, "Federated learning with local differential privacy: Trade-offs between privacy, utility, and communication," in *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 2650–2654. https://doi.org/10.1109/ICASSP39728.2021.9413764

[6] Y. Li, S. Wang, C. Y. Chi, and T. Q. Quek, "Differentially private federated clustering over non-IID data," *IEEE Internet of Things Journal*, vol. 11, no. 4, pp. 6705–6721, 2023. https://doi.org/10.1109/JIOT.2023.3312852

[7] Y. B. Zikria, M. K. Afzal, and S. W. Kim, "Internet of Multimedia Things (IoMT): Opportunities, challenges and solutions," *Sensors*, vol. 20, no. 8, p. 2334, 2020. https://doi.org/10.3390/s20082334

[8] A. Barnawi, P. Chhikara, R. Tekchandani, N. Kumar, and B. Alzahrani, "A differentially privacy assisted federated learning scheme to preserve data privacy for IoMT applications," *IEEE Transactions on Network and Service Management*, vol. 21, no. 4, pp. 4686–4700, 2024. https://doi.org/10.1109/TNSM.2024.3393969

[9] J. Fu, Z. Chen, and X. Han, "Adap DP-FL: Differentially private federated learning with adaptive noise," in *2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2022, pp. 656–663. https://doi.org/10.1109/TrustCom56396.2022.00094

[10] C. Singh, R. Mishra, H. P. Gupta, and G. Banga, "A federated learning-based patient monitoring system in Internet of Medical Things," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 4, pp. 1622–1628, 2022. https://doi.org/10.1109/TCSS.2022.3228965

[11] R. Ramani, A. R. Mary, S. E. Raja, and D. A. Shunmugam, "Optimized data management and secured federated learning in the Internet of Medical Things (IoMT) with blockchain technology," *Biomedical Signal Processing and Control*, vol. 93, p. 106213, 2024. https://doi.org/10.1016/j.bspc.2024.106213

[12] A. K. Nair, J. Sahoo, and E. D. Raj, "Privacy preserving federated learning framework for IoMT based big data analysis using edge computing," *Computer Standards & Interfaces*, vol. 86, p. 103720, 2023. https://doi.org/10.1016/j.csi.2023.103720

[13] M. V. Luzón *et al.*, "A tutorial on federated learning from theory to practice: Foundations, software frameworks, exemplary use cases, and selected trends," *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 4, pp. 824–850, 2024. https://doi.org/10.1109/JAS.2024.124215

[14] S. A. Farooqi, A. Abd Rahman, and A. Saad, "Differential privacy based federated learning techniques in IoMT: A review," in *2024 18th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, 2024, pp. 1–7. https://doi.org/10.1109/IMCOM60618.2024.10418361

[15] M. Iqbal, A. Tariq, M. Adnan, I. U. Din, and T. Qayyum, "FL-ODP: An optimized differential privacy enabled privacy preserving federated learning," *IEEE Access*, vol. 11, pp. 116674–116683, 2023. https://doi.org/10.1109/ACCESS.2023.3325396

[16] B. Zhang, Y. Mao, Z. Tu, X. He, P. Ping, and J. Wu, "Optimizing privacy-accuracy trade-off in DP-FL via significant gradient perturbation," in *2023 19th International Conference on Mobility, Sensing and Networking (MSN)*, 2023, pp. 423–430. https://doi.org/10.1109/MSN60784.2023.00068

[17] F. Z. Errounda and Y. Liu, "Adaptive differential privacy in vertical federated learning for mobility forecasting," *Future Generation Computer Systems*, vol. 149, pp. 531–546, 2023. https://doi.org/10.1016/j.future.2023.07.033

[18] L. Chen, D. Yue, X. Ding, Z. Wang, K. K. R. Choo, and H. Jin, "Differentially private deep learning with dynamic privacy budget allocation and adaptive optimization," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4422–4435, 2023. https://doi.org/10.1109/TIFS.2023.3293961

[19] Y. Mao, C. Li, Z. Wang, Z. Tu, and P. Ping, "Differential privacy in federated dynamic gradient clipping based on gradient norm," in *Algorithms and Architectures for Parallel Processing, ICA3PP 2023*, in Lecture Notes in Computer Science, Z. Tari, K. Li, and H. Wu, Eds., Springer, Singapore, vol. 14490, 2023, pp. 24–41. https://doi.org/10.1007/978-981-97-0859-8_2

[20] A. Bkakria, A. Tasidou, N. Cuppens-Boulahia, F. Cuppens, F. Bouattour, and F. Ben Fredj, "Optimal distribution of privacy budget in differential privacy," in *Risks and Security of Internet and Systems, CRiSIS 2018*, in Lecture Notes in Computer Science, A. Zemmari, M. Mosbah, N. Cuppens-Boulahia, and F. Cuppens, Eds., Springer, Cham, vol. 11391, 2019, pp. 222–236. https://doi.org/10.1007/978-3-030-12143-3_18

[21] J. Li, T. Chen, and S. Teng, "A comprehensive survey on client selection strategies in federated learning," *Computer Networks*, vol. 251, p. 110663, 2024. https://doi.org/10.1016/j.comnet.2024.110663

[22] H. Fereidooni *et al.*, "SAFELearn: Secure aggregation for private federated learning," in *2021 IEEE Security and Privacy Workshops (SPW)*, 2021, pp. 56–62. https://doi.org/10.1109/SPW53761.2021.00017

[23] X. Li, J. Ning, G. S. Poh, L. Y. Zhang, X. Yin, and T. Zhang, "Fluent: Round-efficient secure aggregation for private federated learning," *arXiv preprint arXiv:2403.06143*, 2024.

[24] A. Yazdinejad, A. Dehghantanha, G. Srivastava, H. Karimipour, and R. M. Parizi, "Hybrid privacy preserving federated learning against irregular users in next-generation Internet of Things," *Journal of Systems Architecture*, vol. 148, p. 103088, 2024. https://doi.org/10.1016/j.sysarc.2024.103088

[25] R. Ahmed, P. K. R. Maddikunta, T. R. Gadekallu, N. K. Alshammari, and F. A. Hendaoui, "Efficient differential privacy enabled federated learning model for detecting COVID-19 disease using chest X-ray images," *Frontiers in Medicine*, vol. 11, p. 1409314, 2024. https://doi.org/10.3389/fmed.2024.1409314

[26] X. Cao *et al.*, "Communication-efficient distributed learning: An overview," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 4, pp. 851–873, 2023. https://doi.org/10.1109/JSAC.2023.3242710

[27] J. Yun, Y. Oh, Y. S. Jeon, and H. V. Poor, "Communication-efficient federated learning over capacity-limited wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 1, pp. 621–637, 2024. https://doi.org/10.1109/TCCN.2024.3419039

[28] G. Gad, E. Gad, Z. M. Fadlullah, M. M. Fouda, and N. Kato, "Communication-efficient and privacy-preserving federated learning via joint knowledge distillation and differential privacy in bandwidth-constrained networks," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 11, pp. 17586–17601, 2024. https://doi.org/10.1109/TVT.2024.3423718

[29] S. C. Messinis, N. E. Protonotarios, and N. Doulamis, "Differentially private client selection and resource allocation in federated learning for medical applications using graph neural networks," *Sensors*, vol. 24, no. 16, p. 5142, 2024. https://doi.org/10.3390/s24165142

[30] I. Mironov, "Rényi differential privacy," in *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, 2017, pp. 263–275. https://doi.org/10.1109/CSF.2017.11

[31] D. N. Sachin, B. Annappa, S. Hegde, C. S. Abhijit, and S. Ambesange, "Fedcure: A heterogeneity-aware personalized federated learning framework for intelligent healthcare applications in IoMT environments," *IEEE Access*, vol. 12, pp. 15867–15883, 2024. https://doi.org/10.1109/ACCESS.2024.3357514

[32] S. H. A. Kazmi, R. Hassan, F. Qamar, K. Nisar, and D. P. Dahnil, "Threat intelligence in IoMTs with federated learning using non-IID data: An experimental analysis," in *2024 IEEE 7th International Symposium on Telecommunication Technologies (ISTT)*, 2024, pp. 120–125. https://doi.org/10.1109/ISTT63363.2024.10750596

[33] A. Imteaj, U. Thakker, S. Wang, J. Li, and M. H. Amini, "A survey on federated learning for resource-constrained IoT devices," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 1–24, 2021. https://doi.org/10.1109/JIOT.2021.3095077

[34] S. C. Messinis, N. E. Protonotarios, and N. Doulamis, "Differentially private client selection and resource allocation in federated learning for medical applications using graph neural networks," *Sensors*, vol. 24, no. 16, p. 5142, 2024. https://doi.org/10.3390/s24165142

[35] L. Shi, J. Shu, W. Zhang, and Y. Liu, "HFL-DP: Hierarchical federated learning with differential privacy," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–7. https://doi.org/10.1109/GLOBECOM46510.2021.9685644

[36] Y. Zhou, Q. Ye, and J. Lv, "Communication-efficient federated learning with compensated overlap-fedavg," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 1, pp. 192–205, 2021. https://doi.org/10.1109/TPDS.2021.3090331

[37] X. Sun, Z. Yuan, X. Kong, L. Xue, L. He, and Y. Lin, "Communication-efficient and privacy-preserving aggregation in federated learning with adaptability," *IEEE Internet of Things Journal*, vol. 11, no. 15, pp. 26430–26443, 2024. https://doi.org/10.1109/JIOT.2024.3396217

[38] M. Yang, H. Cheng, F. Chen, X. Liu, M. Wang, and X. Li, "Model poisoning attack in differential privacy-based federated learning," *Information Sciences*, vol. 630, pp. 158–172, 2023. https://doi.org/10.1016/j.ins.2023.02.025

[39] F. N. Al-Wesabi *et al.*, "Pelican optimization algorithm with federated learning driven attack detection model in internet of things environment," *Future Generation Computer Systems*, vol. 148, pp. 118–127, 2023. https://doi.org/10.1016/j.future.2023.05.029

[40] A. Yazdinejad, A. Dehghantanha, H. Karimipour, G. Srivastava, and R. M. Parizi, "A robust privacy-preserving federated learning model against model poisoning attacks," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 6693–6708, 2024. https://doi.org/10.1109/TIFS.2024.3420126

[41] F. Liang, Z. Zhang, H. Lu, V. Leung, Y. Guo, and X. Hu, "Communication-efficient large-scale distributed deep learning: A comprehensive survey," *arXiv preprint arXiv: 2404.06114*, 2024.

[42] H. M. Kwan and S. Song, "FedSDD: Scalable and diversity-enhanced distillation for model aggregation in federated learning," *arXiv preprint arXiv:2312.17029*, 2023.

[43] Y. Yuan *et al.*, "Distributed learning for large-scale models at edge with privacy protection," *IEEE Transactions on Computers*, vol. 73, no. 4, pp. 1060–1070, 2024. https://doi.org/10.1109/TC.2024.3352814

[44] Z. Chen, G. Liao, Q. Ma, and X. Chen, "Adaptive privacy budget allocation in federated learning: A multi-agent reinforcement learning approach, in *ICC 2024-IEEE International Conference on Communications*, 2024, pp. 5166–5171. https://doi.org/10.1109/ICC51166.2024.10622685

[45] L. Chen, D. Yue, X. Ding, Z. Wang, K. K. R. Choo, and H. Jin, "Differentially private deep learning with dynamic privacy budget allocation and adaptive optimization," *IEEE Transactions on Information Forensics and Security,* vol. 18, pp. 4422–4435, 2023. https://doi.org/10.1109/TIFS.2023.3293961

[46] Y. Mao, C. Li, Z. Wang, Z. Tu, and P. Ping, "Differential privacy in federated dynamic gradient clipping based on gradient norm," in *Algorithms and Architectures for Parallel Processing, ICA3PP 2023,* in Lectures Notes in Computer Science, Z. Tari, K. Li, and H. Wu, Eds., Springer, Singapore, vol. 14490, 2023, pp. 24–41. https://doi.org/10.1007/978-981-97-0859-8_2

[47] J. Fu, Z. Chen, and X. Han, "Adap DP-FL: Differentially private federated learning with adaptive noise," in *2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2022, pp. 656–663. https://doi.org/10.1109/TrustCom56396.2022.00094

[48] K. Wei *et al.*, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020. https://doi.org/10.1109/TIFS.2020.2988575

[49] H. M. F. Noman, K. Dimyati, K. A. Noordin, E. Hanafi, and A. Abdrabou, "FeDRL-D2D: Federated deep reinforcement learning-empowered resource allocation scheme for energy efficiency maximization in D2D-Assisted 6G networks," *IEEE Access*, vol. 12, pp. 109775–109792, 2024. https://doi.org/10.1109/ACCESS.2024.3434619

[50] Y. Zhang, C. Xu, H. H. Yang, X. Wang, and T. Q. Quek, "DPP-based client selection for federated learning with non-IID data," in *ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10095604

[51] Y. Xu *et al.*, "Enhancing decentralized federated learning with model pruning and adaptive communication," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 1, pp. 70–84, 2024. https://doi.org/10.1109/TII.2024.3424497

[52] P. R. Ovi, E. Dey, N. Roy, and A. Gangopadhyay, "Mixed quantization enabled federated learning to tackle gradient inversion attacks," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2023, pp. 5046–5054. https://doi.org/10.1109/CVPRW59228.2023.00533

[53] B. Niu, Y. Chen, B. Wang, Z. Wang, F. Li, and J. Cao, "AdaPDP: Adaptive personalized differential privacy," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*, 2021, pp. 1–10. https://doi.org/10.1109/INFOCOM42981.2021.9488825

[54] Z. Liu, W. Wang, H. Liang, and Y. Yuan, "Enhancing data utility in personalized differential privacy: A fine-grained processing approach," in *Data Security and Privacy Protection, DSPP 2024*, in Lecture Notes in Computer Science, X. Chen, X. Huang, and M. Yung, Eds., Springer, Singapore, vol. 15216, 2024, pp. 47–66. https://doi.org/10.1007/978-981-97-8546-9_3

[55] J. Zhang, W. Yang, Y. Zhang, H. Zheng, and T. Zhang, "DPAdaMod_AGC: Adaptive gradient clipping-based differential privacy," in *2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2024, pp. 950–955. https://doi.org/10.1109/CSCWD61410.2024.10580740

[56] G. Lin *et al.*, "Understanding adaptive gradient clipping in DP-SGD, empirically," *International Journal of Intelligent Systems,* vol. 37, no. 11, pp. 9674–9700, 2022. https://doi.org/10.1002/int.23001

[57] J. Fu, Z. Chen, and X. Han, "Adap DP-FL: Differentially private federated learning with adaptive noise," in *2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2022, pp. 656–663. https://doi.org/10.1109/TrustCom56396.2022.00094

[58] M. J. Mia and M. H. Amini, "QuanCrypt-FL: Quantized homomorphic encryption with pruning for secure federated learning," *arXiv preprint arXiv:2411.05260*, 2024.

[59] X. Shen, H. Jiang, Y. Chen, B. Wang, and L. Gao, "PLDP-FL: Federated learning with personalized local differential privacy," *Entropy*, vol. 25, no. 3, p. 485, 2023. https://doi.org/10.3390/e25030485

[60] V. Rattanawiboonsom, H. Sikandar, U. Thatsaringkharnsakun, and N. Khan, "The role of mobile technologies in tracking cyberbullying trends and social adaptation among teenagers," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 1, pp. 171–186, 2025. https://doi.org/10.3991/ijim.v19i01.52747

[61] V. Rattanawiboonsom and N. Khan, "Blockchain technology in mobile payments: A systematic review of security enhancements in mobile commerce," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 21, pp. 134–148, 2024. https://doi.org/10.3991/ijim.v18i21.52099

[62] N. Khan, M. I. Qureshi, M. Falahat, H. Sikandar, and R. Bt Sham, "Navigating the renewable energy transition: A systematic review of economic and policy strategies for grid integration, stability, and viability," *International Journal of Energy Economics and Policy*, vol. 15, no. 4, pp. 709–723, 2025. https://doi.org/10.32479/ijeep.20348

7 AUTHORS

Shaista Ashraf Farooqi is with the Asia e University (AeU), Wisma Subang Jaya, Jalan SS 15/4, Subang Jaya, Malaysia (E-mail: shaista@safarooqi.com, c70101220003@aeu.edu.my).

Aedah Abd Rahman is with the Asia e University (AeU), Wisma Subang Jaya, Jalan SS 15/4, Subang Jaya, Malaysia (E-mail: aedah.abdrahman@aeu.edu.my).

Amna Saad is with the Universiti Kuala Lumpur, Malaysian Institute of Information Technology, 1016 Jalan Sultan Ismail, 50250 Kuala Lumpur, Malaysia (E-mail: amna@unikl.edu.my).