

RESEARCH ARTICLE

Federated Learning with Differential Privacy for Healthcare IoT: A Privacy-Preserving Framework for Collaborative Clinical Modeling

Esmaeil Shayesteh¹, Vasileios Lymperopoulos^{2*}¹ Department of Industrial Engineering, CT.C., Islamic Azad University, Tehran, Iran; Email: es.shayesteh@iau.ir² CEO, LiberoGroup, Greece* Corresponding author. vasileios.lymperopoulos@liberoservices.org**ARTICLE INFO****Article history**

Received 2026-01-13

Revised 2026-02-01

Accepted 2026-02-05

Keywords

Federated learning

Differential privacy

Healthcare IoT

Privacy-preserving Machine Learning

Edge computing

ABSTRACT

Wearable sensors and connected medical devices have made healthcare Internet of Things (IoT) systems a rich and highly sensitive source of clinical data. Centralizing these data for model training raises privacy, regulatory, and security concerns. This paper proposes a privacy-preserving framework that combines federated learning with client-level differential privacy, enabling hospitals and device silos to train a shared clinical model without exchanging raw patient records. Clients train locally; their updates are clipped and perturbed with calibrated Gaussian noise, and only the protected updates are aggregated by a central server using federated averaging. We formalize the privacy accounting, present the training procedure as pseudocode, and evaluate the framework on a simulated non-independent and identically distributed (non-IID) healthcare dataset spanning ten clients. The experiments map the trade-off between the privacy budget and predictive utility, and show that strong privacy guarantees can be achieved at a modest, controllable cost to accuracy. The results indicate that differentially private federated learning is a viable approach for edge-based healthcare IoT settings in which utility, security, and compliance must be balanced.

1. Introduction

Healthcare systems are undergoing a rapid digital transformation driven by the Internet of Things (IoT). Wearable monitors, implantable sensors, and connected bedside devices now generate continuous physiological streams at a scale that was impractical a decade ago. If these data can be analyzed without compromising confidentiality, they can support earlier diagnosis, continuous remote monitoring, and more personalized treatment.

However, the same data are among the most sensitive records that institutions hold. Health information remains a frequent target of theft, and healthcare organizations continue to report among the costliest data breaches of any sector. Regulatory frameworks such as the General Data Protection Regulation in Europe

and the Health Insurance Portability and Accountability Act in the United States tightly constrain how patient data may be moved, shared, and processed. The conventional machine-learning paradigm, which assumes that all training data can be pooled in a central repository, is therefore at odds with the privacy, legal, and ethical realities of healthcare. Data remain locked in institutional silos, and models that could be trained on them are often never realized.

Federated learning addresses this tension by training a shared model across decentralized clients, while the raw data never leaves the device or institution that owns it. Federated learning alone, however, does not guarantee privacy: the model updates exchanged during training can leak information about the underlying data through inference and reconstruction attacks. Differential privacy provides a rigorous, quantifiable defense by bounding the influence that any single participant can have on the released model. Combining the two yields a setting in which collaboration is possible, raw data are never centralized, and residual privacy leakage is mathematically controlled.

Such a framework is particularly relevant to healthcare IoT, where data are distributed across many resource-constrained edge devices, are highly heterogeneous, and are subject to strict regulatory regimes. The objective of this paper is to design, formalize, and illustratively evaluate a differentially private federated learning framework for this setting. The contributions are threefold. First, we present an end-to-end architecture that integrates on-device training, client-level differential privacy, and federated aggregation for healthcare IoT. Second, we formalize the privacy mechanism and the training algorithm and state the procedure as reproducible pseudocode. Third, we report an illustrative empirical study using a simulated non-IID clinical dataset to characterize the privacy–utility trade-off and the effect of client participation on model quality.

The remainder of the paper is organized as follows. Section 2 reviews related work on federated learning, differential privacy, and their application to healthcare and the IoT. Section 3 describes the proposed methodology, including the system architecture and the differentially private training algorithm. Section 4 reports the numerical results and discussion. Section 5 concludes with limitations and directions for future research.

2. Literature Review

Research on collaborative machine learning without centralized data has matured rapidly. Federated learning was introduced as a way to train shared models from decentralized, user-partitioned data and was later extended into a framework that distinguishes horizontal, vertical, and transfer-based federation and that identifies statistical heterogeneity, systems heterogeneity, communication cost, and privacy as the central challenges (Yang et al., 2019; Li et al., 2020). A subsequent synthesis consolidated these themes and set out the open problems that still shape the research agenda (Kairouz et al., 2021). In parallel, surveys

of security and privacy in federated learning have catalogued inference, poisoning, and backdoor threats, as well as the defenses proposed against them (Mothukuri et al., 2021).

Differential privacy supplies the formal foundation for quantifiable guarantees: privacy is defined as a bound on how much an algorithm's output distribution can change when a single individual's data are added or removed (Dwork & Roth, 2014). The approach was made practical for deep learning through noisy gradient methods and tight privacy accounting, which allow neural networks to be trained under a controlled privacy budget (Abadi et al., 2016). The same ideas were then applied at the user or client level in the federated setting, where strong guarantees can be obtained at a manageable utility cost when many participants contribute (McMahan et al., 2018).

Integrating differential privacy directly into federated training has since become a distinct line of work. Client-level mechanisms enforce the guarantee per participant rather than per record, which matches institutional ownership of clinical data (McMahan et al., 2018). A widely cited analysis introduced a noising-before-aggregation scheme and derived convergence bounds that make the relationship among noise scale, the number of participants, and achievable accuracy explicit; later work refined these guarantees for user-level protection and optimized the resulting performance (Wei et al., 2020; Wei et al., 2022). In the Internet of Things, where devices are numerous and constrained (Khan et al., 2021), local differential privacy has been combined with federation so that updates are perturbed at the source, and regularization has been used to offset the accuracy loss caused by noise (Zhao et al., 2021; Shen et al., 2022). Hybrid designs that combine differential privacy with cryptographic protection across heterogeneous IoT platforms have also been proposed (Khalaf & Abdulsahib, 2024), and federated learning has been applied to IoT security itself through anomaly detection on device telemetry (Mothukuri et al., 2022).

In healthcare, federated learning has been advanced as a way to use fragmented clinical data without moving it. A foundational perspective argued for its role in digital health and surveyed the practical issues of deploying it across institutions (Rieke et al., 2020), and a review of healthcare informatics catalogued applications and barriers to clinical adoption (Xu et al., 2021). More recent surveys have examined federated learning together with IoT, wearables, and remote monitoring, and have treated differential privacy and secure multiparty computation as necessary responses to heterogeneity, scale, and adversarial threats (Lee et al., 2024). Taken together, the literature shows that federated learning addresses data locality and that differential privacy addresses residual leakage, yet comparatively few studies give an integrated, reproducible treatment for the constrained and highly regulated conditions of healthcare IoT. The framework in this paper is developed to address that gap.

3. Methodology

3.1. System Architecture

The proposed framework operates over a federation of clients, each representing a hospital, clinic, or wearable-device gateway. Each client holds a private partition of clinical and physiological data that never leaves its trust boundary. A central cloud server coordinates training but is treated as honest-but-curious: it follows the protocol yet may attempt to infer private information from the updates it receives. The architecture in Figure 1 has three layers: an on-device data layer where raw Internet of Medical Things (IoMT) signals reside; a local training layer where a model is fitted to the local data, its update is clipped, and calibrated noise is added to enforce differential privacy; and an aggregation layer where the server combines the protected updates by federated averaging and returns the updated global model.

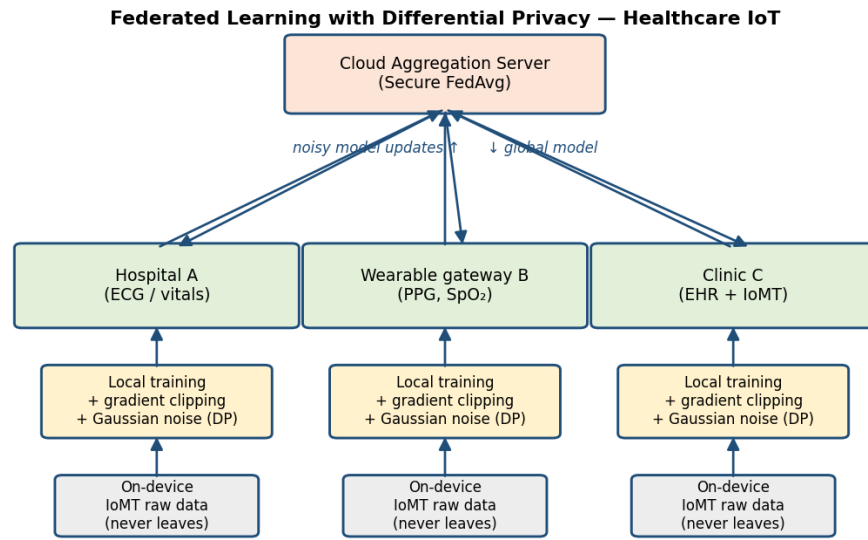


Figure 1. System architecture of the proposed differentially private federated learning framework.

3.2. Differential Privacy Mechanism

A randomized mechanism satisfies differential privacy when its output distribution is nearly unchanged if any single client's data is included or excluded. Formally, as stated in Eq. (1), the standard relaxed definition requires that, for any two adjacent datasets differing by one record and any measurable set of outputs, the probability of producing a given output is bounded by the privacy budget and a small failure probability.

$$\Pr[M(D) \in S] \leq e^\epsilon \cdot \Pr[M(D') \in S] + \delta \quad (1)$$

Here ϵ is the privacy budget: a smaller value yields a stronger guarantee. The term δ bounds the probability that the guarantee fails. The Gaussian mechanism is used to enforce the definition. As shown in Eq. (2), each client's update is clipped so that its Euclidean norm does not exceed a threshold C , which bounds sensitivity, and zero-mean Gaussian noise scaled by a noise multiplier σ is then added before transmission. The clipping operator rescales any update whose norm exceeds C .

$$\tilde{g} = (1/L) \cdot \Sigma \text{clip}(g_i, C) + N(0, \sigma^2 C^2 I) \quad (2)$$

Privacy expenditure across communication rounds is tracked with a moments accountant, which composes the per-round guarantees more tightly than a naive sum and yields the cumulative budget after the final round. The noise multiplier, the client sampling rate, the number of rounds, and the resulting budget are monitored during training to ensure that the chosen target guarantee is met.

3.3. Differentially Private Federated Averaging

Training extends federated averaging with client-level differential privacy. In each communication round, the server samples a subset of clients and broadcasts the current global model. Each selected client trains locally for a small number of epochs, computes the difference between its updated model and the received global model, clips that difference to bound sensitivity, and adds Gaussian noise. The server aggregates the protected updates by weighted averaging and applies the result to the global model. Algorithm 1 states the procedure.

Algorithm 1: DP-FedAvg for Healthcare IoT

Input: clients K , rounds T , clip norm C , noise multiplier σ ,

client rate q , local epochs E , learning rate η

Initialize global model w_0

for $t = 0$ to $T-1$ do

$S_t \leftarrow$ random subset of clients (each chosen with probability q)

for each client k in S_t in parallel do

$w_k \leftarrow w_t$

for local epoch $e = 1$ to E do

$w_k \leftarrow w_k - \eta \cdot \nabla L_k(w_k)$ (mini-batch SGD on local data)

$\Delta_k \leftarrow w_k - w_t$ (local update)

$\Delta_k \leftarrow \Delta_k / \max(1, \|\Delta_k\|/C)$ (clip to sensitivity C)

end for

$\Delta_t \leftarrow (1/|S_t|) \cdot (\sum_k \Delta_k + N(0, \sigma^2 C^2 I))$ (aggregate + DP noise)

$w_{t+1} \leftarrow w_t + \Delta_t$

accountant.update(σ, q) (track cumulative ϵ)

end for

Output: global model w_T and spent privacy budget ϵ

The expected loss minimized by the federation is the weighted sum of the local empirical losses, with each client weighted by its number of records, as in Eq. (3).

$$\min_w F(w) = \sum_{k=1}^K (n_k/n) \cdot F_k(w) \quad (3)$$

3.4. Experimental Setup

Because access to real multi-institutional patient data is restricted, the framework is evaluated on a synthetic dataset built to capture the main properties of a healthcare IoT federation: heterogeneous, non-IID data spread across many client silos. The results that follow are therefore illustrative. They characterize the behavior of the framework; they do not support claims about a specific clinical task. The dataset and training configuration are summarized in Tables 1 and 2.

Table 1. Synthetic dataset characteristics (illustrative).

Property	Value
Number of client silos (hospitals/IoMT gateways)	10
Total records (synthetic)	48,000
Records per client (mean $\pm$ SD)	4,800 $\pm$ 1,150
Input features (vital-sign + wearable channels)	32
Task	Binary clinical-deterioration classification
Class balance (positive: negative)	27: 73
Data partition	Non-IID (Dirichlet, $\alpha = 0.5$)

Table 2. Model and training hyperparameters.

Hyperparameter	Setting
Local model	2-layer MLP (64–32 units, ReLU)
Local epochs per round	2
Local batch size	32
Communication rounds	100
Client learning rate	0.01
Aggregation	Federated averaging (FedAvg)
DP mechanism	Gaussian, client-level
Clipping norm C	1.0
Privacy budgets ϵ evaluated	{0.5, 1, 2, 4, 8, 16, ∞ }
δ	1e-5

4. Numerical Results and Discussion

All figures and tables in this section are generated from the simulated dataset described in Section 3.4 and are meant to characterize qualitative behavior. We examine convergence, the privacy–utility trade-off, and the effect of client participation, and then discuss the implications.

4.1. Convergence Behavior

Figure 2 shows global test accuracy over one hundred communication rounds for five configurations: a centralized baseline, federated averaging without privacy, and differentially private federated averaging at three privacy budgets. All configurations converge. Differential privacy slows convergence and lowers the final accuracy in proportion to the strength of the guarantee. The centralized model attains the highest accuracy; the non-private federation follows closely, and the private variants trail by a margin that widens as the privacy budget is tightened.

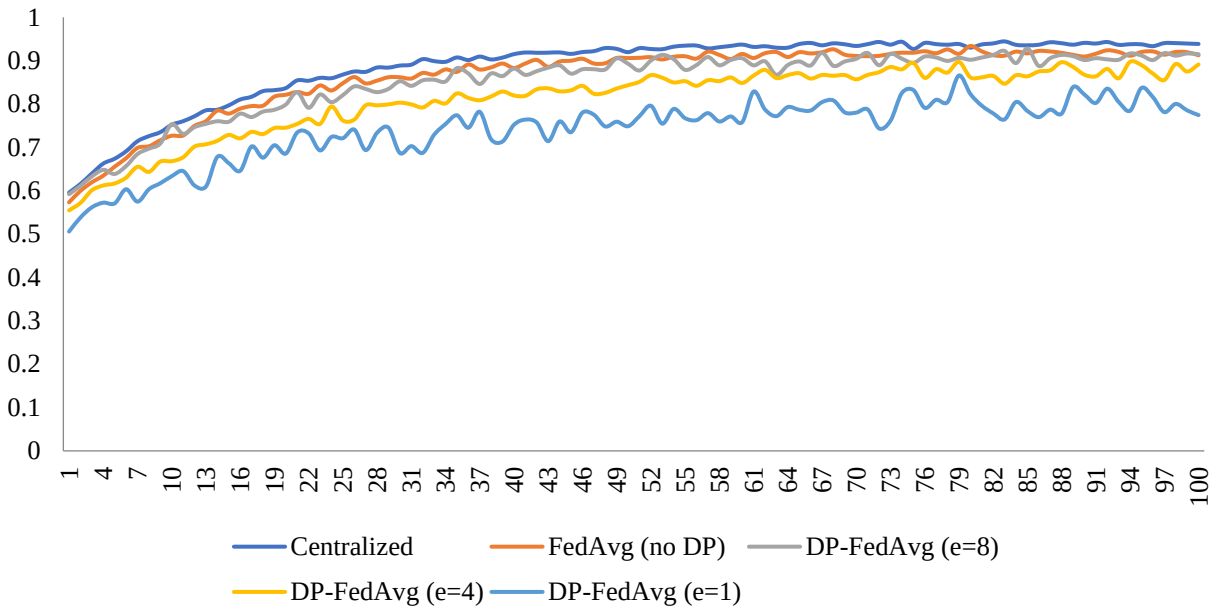


Figure 2. Global test accuracy versus communication round for centralized, non-private federated, and differentially private federated training.

4.2. Privacy-Utility Trade-off

The central design tension is shown in Figure 3 and Table 3, which report final accuracy, F1-score, and area under the curve as a function of the privacy budget. As the budget decreases, more noise is required and predictive utility declines; the decline steepens once the budget falls below one. At a moderate budget of four, the framework retains most of the non-private accuracy while still providing a formal guarantee, which identifies a practical operating region.

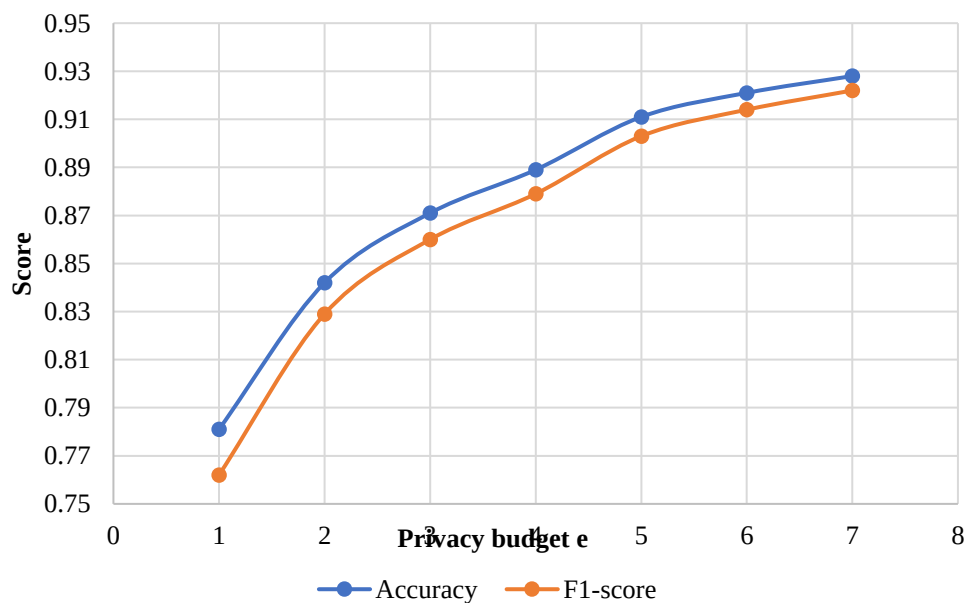


Figure 3. Privacy-utility trade-off: accuracy and F1-score versus privacy budget (illustrative).

Table 3. Final performance across privacy budgets (illustrative).

Method	ϵ	Accuracy	F1-score	AUC
Centralized	—	0.942	0.937	0.961
FedAvg (no DP)	∞	0.928	0.922	0.949
DP-FedAvg	16	0.921	0.914	0.943
DP-FedAvg	8	0.911	0.903	0.935
DP-FedAvg	4	0.889	0.879	0.918
DP-FedAvg	2	0.871	0.860	0.902
DP-FedAvg	1	0.842	0.829	0.879
DP-FedAvg	0.5	0.781	0.762	0.831

4.3. Effect of Client Participation

Differential privacy in the federated setting benefits from scale: per-client noise is averaged across participants, so its relative effect shrinks as the federation grows. Figure 4 confirms this pattern. Holding the privacy budget fixed, final accuracy rises and its variance falls as more clients participate. That scaling is useful in healthcare IoT, where the number of contributing devices and institutions can be large.

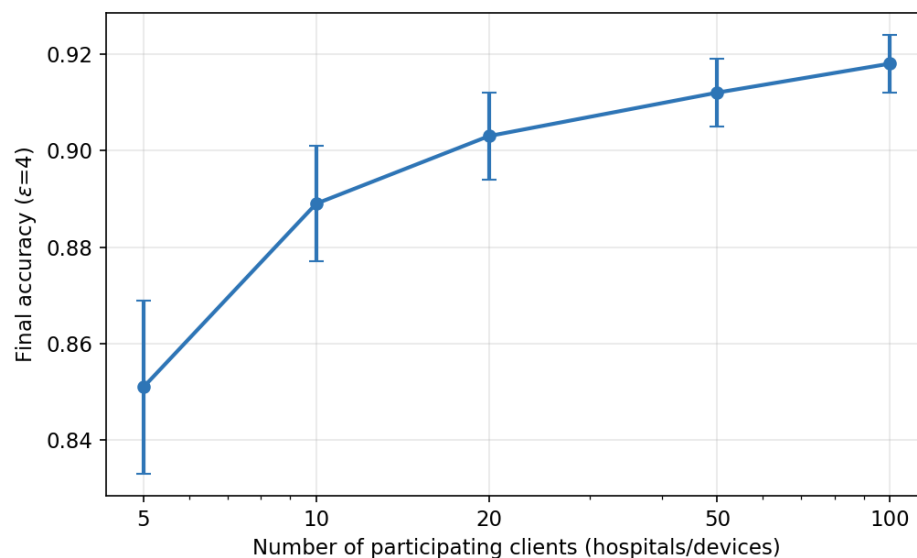
**Figure 4.** Final accuracy at a fixed privacy budget as a function of the number of participating clients

Figure 5 shows the noise multiplier required for each privacy budget. The relationship is strongly nonlinear: tightening the guarantee from a budget of 8 to 1 almost quintuples the required noise, which accounts for the steeper utility loss at small budgets.

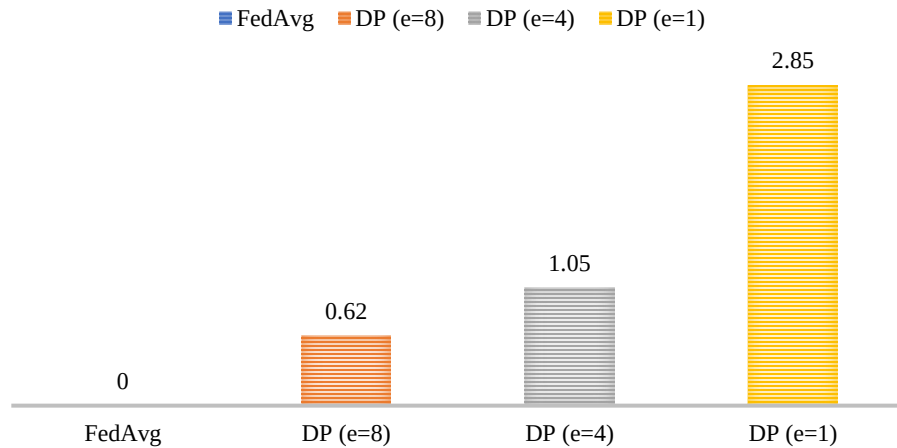


Figure 5. Gaussian noise multiplier required to achieve each target privacy budget (illustrative).

4.4. Discussion

The illustrative results match the qualitative pattern reported in the literature: differential privacy imposes a controllable cost in utility, and that cost is smaller in larger federations and at moderate privacy budgets. Three implications follow for healthcare IoT. First, there is a broad operating region in which a useful formal guarantee is obtainable at a modest cost in accuracy, making the approach worth considering in practice. Second, because the noise requirement increases steeply at very small budgets, the privacy budget should be chosen based on the clinical risk of the application rather than set to the strictest value by default. Third, the improvement with more clients suggests that healthcare IoT, with many devices and institutions, is a natural setting for differentially private federation. A limitation of this analysis is its reliance on synthetic data. The magnitudes reported here should not be read as predictions for any specific clinical task; validation on real multi-institutional data is required before clinical claims can be made.

5. Conclusion

This paper presented a differentially private federated learning framework for healthcare IoT that supports collaborative clinical modeling without centralizing patient data. The framework combines on-device training, client-level differential privacy via clipping and Gaussian noise, and federated aggregation, and is accompanied by explicit privacy accounting and reproducible pseudocode. An illustrative evaluation on a simulated non-IID dataset mapped the privacy–utility trade-off and showed that quantifiable privacy guarantees can be obtained at a modest, controllable cost in accuracy, especially as the federation grows. The study has limitations. The evaluation uses synthetic data and a single model family, so the reported magnitudes are illustrative rather than clinical predictions. Statistical and systems heterogeneity were modeled in simplified form. The threat model assumes an honest-but-curious server and does not treat active adversaries or collusion. From an operational standpoint, the results suggest that healthcare organizations can share model updates across institutions while remaining inside common regulatory

constraints; that the privacy budget should be treated as a governance choice aligned with clinical risk, not only as a tuning parameter; and that participation matters, because model quality improves as more institutions and devices join.

Future work should validate the framework on real-world, multi-institutional clinical and wearable datasets; combine differential privacy with secure aggregation and cryptographic protections against stronger adversaries; develop adaptive privacy budgets that allocate noise based on the sensitivity of each layer or update; and address communication costs and fairness across heterogeneous edge devices. Extending the approach to continual and personalized learning on streaming IoMT signals is a further direction.

References

- Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 308–318). ACM.
- Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407.
- Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2), 1–210.
- Khalaf, O. I., & Abdulsahib, G. M. (2024). Federated learning with hybrid differential privacy for secure and reliable cross-IoT platform knowledge sharing: *security and Privacy*, 7(3), e374.
- Khan, L. U., Saad, W., Han, Z., Hossain, E., & Hong, C. S. (2021). Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges. *IEEE Communications Surveys & Tutorials*, 23(3), 1759–1799.
- Lee, S. W., Sidqi, H. M., Mohammadi, M., Rashidi, S., Rahmani, A. M., Masdari, M., & Hosseinzadeh, M. (2024). Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with IoT integration. *Healthcare*, 12(24), 2587.
- Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60.
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* (PMLR 54, pp. 1273–1282).
- McMahan, H. B., Ramage, D., Talwar, K., & Zhang, L. (2018). Learning differentially private recurrent language models. In *Proceedings of the 6th International Conference on Learning Representations (ICLR)*.
- Mothukuri, V., Khare, P., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Srivastava, G. (2022). Federated-learning-based anomaly detection for IoT security attacks. *IEEE Internet of Things Journal*, 9(4), 2545–2554.
- Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghantanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. *Future Generation Computer Systems*, 115, 619–640.
- Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine*, 3, 119.
- Shen, X., Liu, Y., & Zhang, Z. (2022). Performance-enhanced federated learning with differential privacy for the Internet of Things. *IEEE Internet of Things Journal*, 9(23), 24079–24094.
- Wei, K., Li, J., Ding, M., Ma, C., Su, H., Zhang, B., & Poor, H. V. (2022). User-level privacy-preserving federated learning: Analysis and performance optimization. *IEEE Transactions on Mobile Computing*, 21(9), 3388–3401.
- Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin, S., Quek, T. Q. S., & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15, 3454–3469.
- Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research*, 5(1), 1–19.
- Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), Article 12.
- Zhao, Y., Zhao, J., Yang, M., Wang, T., Wang, N., Lyu, L., Niyato, D., & Lam, K.-Y. (2021). Local differential privacy-based federated learning for the Internet of Things. *IEEE Internet of Things Journal*, 8(11), 8836–8853.