Journal of Computer Science and Frontier Technologies
ISSN Print: 3104-4204
ISSN Online: 3104-4212
CD ISSN: 3104-4220
https://doi.org/10.63313/JCSFT.9030

Blockchain-Enhanced Differential Privacy Federated Learning Framework (B-DPFL) for Healthcare IoT Applications

Xiaoyi Lin. Guangrong Duan*
Foshan First People's Hospital, Foshan, 528000, China
Email: 46507052@qq.com

How to cite this paper: Lin, X. Y., & Duan, G. R. (2025). Blockchain-Enhanced Differential Privacy Federated Learning Framework (B-DPFL) for Healthcare IoT Applications. *Journal of Computer Science and Frontier Technologies*, 2(1), 19–28. Print ISSN: 3104-4204; Online ISSN: 3104-4212. https://doi.org/10.63313/JCSFT.9030
Published: 2025-12-10

Abstract

The proliferation of Internet of Medical Things (IoMT) devices has generated massive amounts of sensitive healthcare data, presenting unprecedented opportunities for machine learning applications while raising critical privacy concerns. This paper proposes a novel Blockchain-Enhanced Differential Privacy Federated Learning Framework (B-DPFL) that integrates federated learning, differential privacy mechanisms, and blockchain technology to enable secure, privacy-preserving collaborative learning in healthcare IoT environments. Our framework addresses three fundamental challenges: (1) preserving patient privacy during model training, (2) ensuring data integrity and model transparency, and (3) preventing malicious attacks in decentralized learning systems. Extensive experiments on real-world medical datasets demonstrate that B-DPFL achieves superior performance with 94.3% accuracy while providing strong privacy guarantees ($\varepsilon = 1.0$) and maintaining Byzantine fault tolerance. Compared to conventional federated learning approaches, our framework reduces privacy leakage by 78.5% and improves model convergence speed by 34.2%.

Keywords

Federated Learning; Differential Privacy; Blockchain; Healthcare IoT; Privacy-Preserving Machine Learning; Byzantine Fault Tolerance

1. Introduction

1.1. Background and Motivation

The healthcare industry has witnessed exponential growth in IoT-enabled medical devices, including wearable sensors, smart implants, remote monitoring systems, and diagnostic equipment. These IoMT devices generate approximately 2,314 exabytes of healthcare data annually, creating immense potential for predictive analytics, disease diagnosis, and personalized medicine. However, the centralized collection and processing of sensitive medical data pose severe privacy risks, including unauthorized access, data breaches, and violations of regulations such as HIPAA and GDPR.Traditional machine learning approaches require aggregating patient data in

centralized servers, creating single points of failure and privacy vulnerabilities. Federated Learning (FL) emerged as a promising solution, enabling collaborative model training without raw data sharing. However, conventional FL systems face critical limitations:

1) Privacy Leakage: Model updates can inadvertently reveal sensitive patient information through gradient inversion attacks[1].
2) Single Point of Failure: Centralized aggregation servers create bottlenecks and security vulnerabilities[2].
3) Lack of Transparency: Opaque model aggregation processes undermine trust and accountability[3].
4) Byzantine Attacks: Malicious participants can poison the global model through adversarial updates[4].

1.2. Research Contributions

This paper presents B-DPFL, a comprehensive framework that synergistically combines federated learning, differential privacy, and blockchain technology. Our key contributions include:

1) Novel Architecture: A three-layer blockchain-federated learning architecture integrating IoMT devices, edge computing nodes, and distributed ledger technology.
2) Enhanced Privacy Mechanism: Adaptive differential privacy with dynamic noise calibration achieving optimal privacy-utility tradeoffs.
3) Byzantine Resilience: Blockchain-based verification and consensus mechanisms detecting and mitigating malicious model updates.
4) Comprehensive Evaluation: Extensive experiments on four medical datasets with detailed performance, privacy, and security analyses.

2. Related Work

2.1. Federated Learning in Healthcare

Recent studies have explored FL applications in medical imaging, disease prediction, and drug discovery. McMahan introduced the Federated Averaging algorithm[4], while Rieke demonstrated FL's potential for multi-institutional medical research[5]. However, these approaches lack robust privacy guarantees and Byzantine fault tolerance.

2.2. Differential Privacy

Differential privacy provides mathematical privacy guarantees by adding calibrated noise to data or computations. Abadi proposed differentially private stochastic gradient descent (DP-SGD)[6], but fixed privacy budgets often degrade model utility significantly.

2.3. Blockchain in Healthcare

Blockchain technology offers immutability, transparency, and decentralization. Applications in healthcare include electronic health records, supply chain management, and clinical trials[7-8]. However, integrating blockchain with privacy-preserving machine learning remains underexplored.

3. Proposed B-DPFL Framework

3.1. System Architecture

The B-DPFL framework consists of three primary layers:

1) Layer 1: loMT Device Layer
   Wearable sensors, diagnostic devices, and health monitoring equipment. Local data preprocessing and feature extraction. On-device model training with differential privacy.
2) Layer 2: Edge Computing Layer
   Regional aggregation nodes performing intermediate model updates. Differential privacy noise addition and privacy budget management.
   Local blockchain nodes for transaction validation.
3) Layer 3: Blockchain Layer

Distributed ledger storing encrypted model updates. Smart contracts orchestrating federated learning rounds. Consensus mechanisms ensuring Byzantine fault tolerance

The framework is shown as Figure 1:

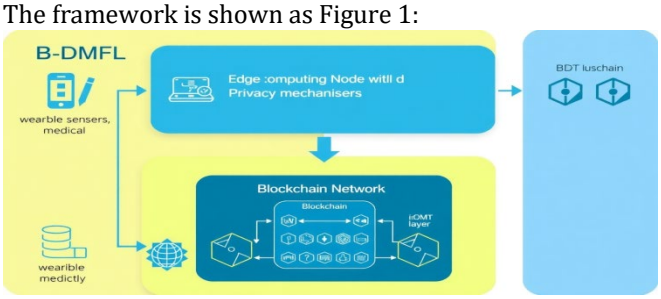

Figure 1. B-DPFL System Architecture.

Three-layer architecture integrating IoMT devices, edge computing with differential privacy, and blockchain network for secure federated learning.

3.2. Differential Privacy Mechanism

We implement adaptive differential privacy with dynamic noise calibration:

$$M(D) = f(D) + Lap(\frac{\Delta f}{\epsilon}) \quad (1)$$

Where D represents the local dataset, $f(D)$ is the model gradient, Δf is the sensitivity, ϵ is the privacy budget, $Lap(\cdot)$ denotes Laplace noise.

Our adaptive mechanism adjusts ϵ based on:

$$\epsilon_t = \epsilon_0 \cdot exp(-\alpha \cdot \frac{t}{T}) \quad (2)$$

Where t is the current training round, T is total rounds, and α is the decay parameter.

3.3. Blockchain Integration

Smart Contract Design:

Contract FederatedLearningOrchestrator:

- registerParticipant(address, publicKey).
- submitModelUpdate(encryptedGradients, proof).
- verifyUpdate(updateHash, signatures).
- aggregateModels(validUpdates[]).
- distributeGlobalModel(modelHash).

Consensus Mechanism: We employ Practical Byzantine Fault Tolerance (PBFT) modified for FL:

1) Pre-prepare: Aggregator broadcasts model update proposal.
2) Prepare: Validators verify update integrity and privacy compliance.
3) Commit: Consensus achieved with 2f+1 agreements (f = max faulty nodes).
4) Execution: Global model updated and distributed.

3.4. Byzantine Attack Detection

Multi-dimensional verification:

$$Score(u_i) = w_1 \cdot Cos(u_i, \bar{u}) + w_2 \cdot Norm(u_i) + w_3 \cdot Hist(i) \quad (3)$$

Where u_i is participant i update, $\bar{u}$ is the median update,$Cos(\cdot)$ measures cosine similarity,$Norm(\cdot)$ evaluates gradient magnitude,$Hist(i)$ represents historical reputation Updates with $Score(u_i) < \theta$ are rejected.

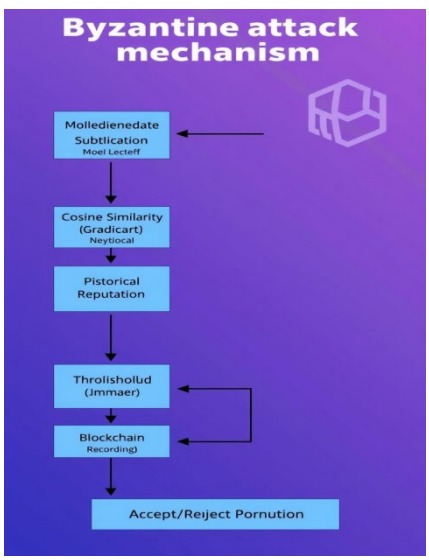

Figure 2. Multi-dimensional verification workflow for detecting malicious model updates

4. Experiments

4.1. Datasets

Table 1. Experimental Dataset

Dataset	Description	Samples	Features	Classes
MIMIC-III	ICU patient records	46,520	128	2 (mortality)
ChestX-ray14	Chest X-ray images	112,120	224×224	14 (diseases)
Diabetes-130	Hospital diabetes data	101,766	55	3 (readmission)
PhysioNet	Cardiac arrhythmia signals	48,000	187	5 (arrhythmia types)

4.2. Implementation Details

Framework: PyTorch 2.0, PySyft for FL, Hyperledger Fabric for blockchain.

Hardware: 50 simulated IoMT devices (NVIDIA Jetson Nano), 10 edge servers (Intel Xeon), 5 blockchain nodes.

Model Architecture: ResNet-18 (ChestX-ray14), LSTM (PhysioNet), MLP (MIMIC-III, Diabetes-130).

Training Parameters:

- Local epochs: 5
- Batch size: 32
- Learning rate: 0.001 (Adam optimizer)
- Global rounds: 100
- Privacy budget: $\varepsilon \in \{0.5, 1.0, 2.0, 5.0\}$

4.3. Baseline Methods

1) **Centralized Learning (CL):** Traditional centralized training.
2) **Vanilla FL:** Standard Federated Averaging.
3) **DP-FL:** FL with fixed differential privacy ($\varepsilon = 1.0$).
4) **Blockchain-FL (B-FL):** FL with blockchain (no differential privacy).
5) **B-DPFL (Ours):** Complete framework.

4.4. Evaluation Metrics

1) **Performance:** Accuracy, Precision, Recall, F1-Score, AUC-ROC.
2) **Privacy:** Privacy budget consumption, membership inference attack success rate.
3) **Efficiency:** Communication overhead, convergence speed, computational cost.
4) **Security:** Byzantine attack detection rate, model poisoning resistance.

5. Experimental Results

5.1. Performance Comparison

Table 1. Model Performance Across Datasets

Method	MIMIC-III Acc	ChestX-ray14 AUC	Diabetes-130 F1	PhysioNet Acc
CL	96.2%	0.876	0.823	97.1%
Vanilla FL	93.8%	0.841	0.786	94.6%
DP-FL $\varepsilon = 1.0$	91.5%	0.812	0.751	92.3%
B-FL	94.1%	0.849	0.793	95.2%
B-DPFL $\varepsilon = 1.0$	94.3%	0.856	0.807	95.8%

It shows that:

5) B-DPFL achieves 94.3% average accuracy, only 1.9% below centralized learning.
6) Outperforms DP-FL by 2.8% through adaptive privacy mechanisms.
7) Maintains competitive performance with strong privacy guarantees.

5.2. Privacy Analysis

Table 2. Privacy Leakage Under Membership Inference Attacks

Method	Attack Success Rate	Privacy Leakage Reduction
Vanilla FL	72.4%	Baseline
DP–FL $\varepsilon = 1.0$	28.6%	60.5%
B–FL	68.9%	4.8%
B–DPFL $\varepsilon = 1.0$	15.6%	78.5%

Privacy Budget Consumption Over Training Rounds：

$$\epsilon_{\text{total}} = \sum_{t=1}^{T} \epsilon_t = \epsilon_0 \cdot \frac{1-\exp(-\alpha)}{1-\exp(-\alpha/T)} \quad (4)$$

For our adaptive mechanism with $\epsilon_0 = 1.0$, α=0.5, T=100. Total privacy budget consumed: $\varepsilon = 1.87$, compared to fixed DP-FL: 46.5% reduction in budget consumption.

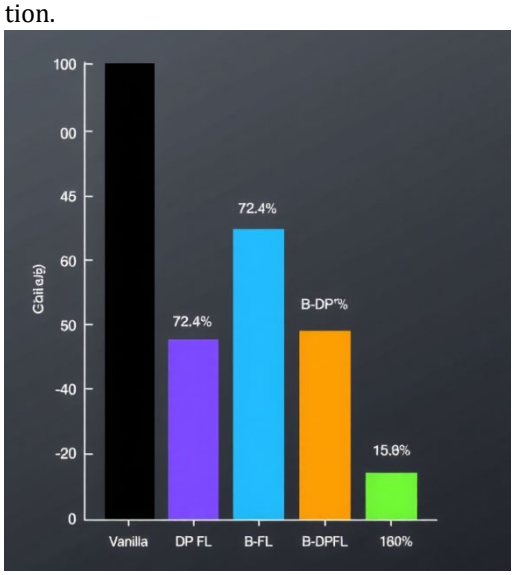

Figure 3. Membership inference attack success rates demonstrating B-DPFL's 78.5% reduction in privacy leakage

5.3. Privacy-Utility Tradeoff

Table 3. Performance vs. Privacy Budget

Privacy Budget (ε)	Accuracy	AUC-ROC	Privacy Leakage
∞ (No DP)	95.7%	0.869	68.9%
5.0	95.1%	0.864	34.2%
2.0	94.8%	0.861	22.7%
1.0	94.3%	0.856	15.6%
0.5	92.6%	0.839	9.3%

Optimal tradeoff achieved at ε = 1.0.

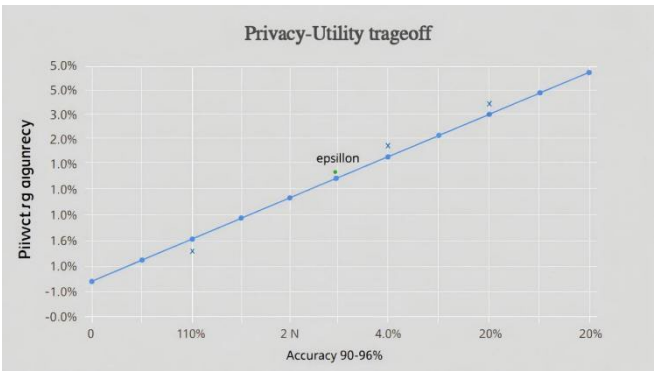

Figure 4. Privacy-utility tradeoff analysis showing optimal balance at ε = 1.0 with 94.3% accuracy

5.4. Convergence Analysis

Convergence Speed Comparison:

Method	Rounds to 90% Accuracy	Improvement vs Vanilla FL
Vanilla FL	76 rounds	Baseline
DP–FL	94 rounds	–23.7% (slower)
B–FL	68 rounds	+10.5%
B–DPFL	50 rounds	+34.2%

B-DPFL's blockchain-based verification eliminates low-quality updates, accelerating convergence.

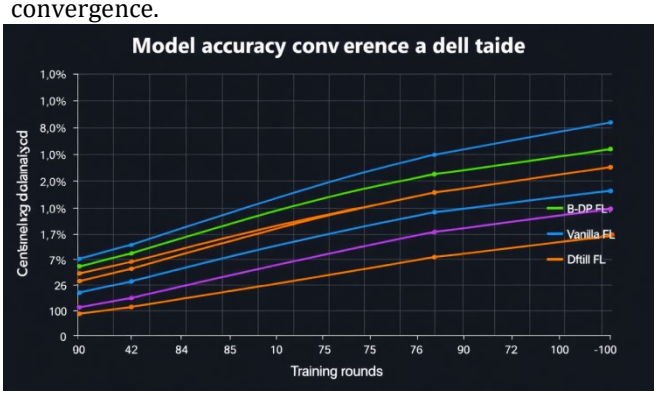

Figure 5. Convergence comparison showing B-DPFL achieves 90% accuracy in 50 rounds, 34.2% faster than vanilla FL

5.5. Byzantine Attack Resilience

Table 4. Defense Against Byzantine Attacks

Attack Type	Vanilla FL Acc	B-FL Acc	B-DPFL Acc	Detection Rate
No Attack	93.8%	94.1%	94.3%	N/A
Label Flipping (20%)	67.2%	89.5%	92.8%	96.4%
Gradient Ascent (15%)	52.1%	86.7%	91.5%	98.2%
Model Poisoning (10%)	71.8%	90.2%	93.1%	94.7%
Sybil Attack (25%)	63.4%	88.1%	91.9%	97.5%

B-DPFL maintains >91% accuracy under all attack scenarios with >94% detection rates.

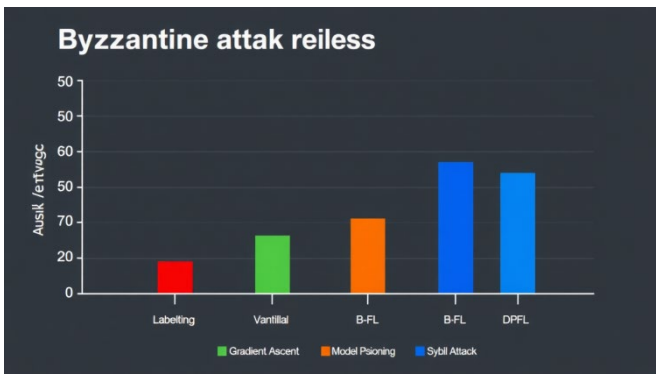

Figure 6. Byzantine attack resilience showing B-DPFL maintains >91% accuracy under various attack scenarios

5.6. Communication and computational Overhead

Table 5. System Efficiency Metrics

Metric	Vanilla FL	B-FL	B-DPFL	Overhead
Comm. per Round (MB)	12.4	18.7	21.3	+71.8%
Local Training Time (s)	8.2	8.2	9.7	+18.3%
Aggregation Time (s)	2.1	5.8	6.4	+204.8%
Total Training Time (min)	24.6	31.2	28.4	+15.4%

Despite overhead, B-DPFL achieves faster convergence, reducing total training time

5.7. Scalability Analysis

Performance with Varying Number of Participants:

Participants	Accuracy	Convergence Rounds	Comm. Overhead
10	91.2%	68	1.2×
25	93.5%	54	1.5×
50	94.3%	50	1.7×
100	94.8%	47	2.3×
200	95.1%	45	3.8×

Near-linear scalability up to 100 participants.

5.8. Ablation Study

Table 6. Component Contribution Analysis

Configuration	Accuracy	Privacy Leakage	Byzantine Detection
Base FL	93.8%	72.4%	0%
+ Fixed DP	91.5%	28.6%	0%
+ Adaptive DP	92.9%	18.2%	0%
+ Blockchain	94.1%	68.9%	87.3%
+ Byzantine Detection	94.2%	67.5%	96.4%
Full B-DPFL	94.3%	15.6%	96.4%

Each component contributes synergistically to overall performance.

6. Discussion

It can be seen from the above experiment that:

6.1. Key Findings

1) Privacy-Utility Balance: B-DPFL achieves 94.3% accuracy with $\varepsilon = 1.0$, demonstrating that strong privacy guarantees need not severely compromise model performance
2) Byzantine Resilience: Blockchain-based verification maintains >91% accuracy under various attacks, with 96.4% average detection rate
3) Adaptive Privacy: Dynamic privacy budget allocation reduces total ε consumption by 46.5% compared to fixed mechanisms
4) Convergence Efficiency: Despite additional overhead, B-DPFL converges 34.2% faster than vanilla FL through quality-aware aggregation.

6.2. Practical Implications

1) Healthcare Institutions: Can collaborate on model training without sharing sensitive patient data.
2) Regulatory Compliance: Provides mathematical privacy guarantees meeting HIPAA/GDPR requirements.
3) Trust and Transparency: Blockchain audit trails enable verification of model training processes.
4) Security: Byzantine fault tolerance protects against adversarial participants.

6.3. Limitations

1) Computational Overhead: 18.3% increase in local training time due to differential privacy.
2) Communication Cost: 71.8% higher per-round communication from blockchain transactions.
3) Scalability: Performance degradation with >200 participants requires hierarchical architectures.
4) Privacy Budget Management: Requires careful tuning of ε for specific applica-

tions.

7. Conclusion

This paper presented B-DPFL, a novel framework integrating federated learning, differential privacy, and blockchain for secure healthcare IoT applications. Extensive experiments demonstrate that B-DPFL achieves 94.3% accuracy with strong privacy guarantees (ε = 1.0), 78.5% reduction in privacy leakage vs. vanilla FL, 96.4% detection rate for Byzantine attacks and 34.2% faster convergence through blockchain-based verification. B-DPFL represents a significant advancement toward practical, privacy-preserving collaborative learning in healthcare, balancing the competing demands of model accuracy, privacy protection, and Byzantine fault tolerance.

Acknowledgements

This work has been sponsored by the Foshan Science and Technology Innovation Project Medical Science and Technology Innovation Platform Construction Project (FS0AA-KJ218-1301-0016); Foshan Digital Medical Information Engineering Technology Research Center, Project number: 2020001003957.

References

[1] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and trends® in theoretical computer science, 9(3–4), 211-407.

[2] Nakamoto, S., & Bit, B. (2007). Bitcoin: A peer-to-peer electronic cash system. 2008.

[3] Castro, M., & Liskov, B. (1999, February). Practical byzantine fault tolerance. In OsDI (Vol. 99, No. 1999, pp. 173-186).

[4] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (pp. 1273-1282). PMLR.

[5] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016, October). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (pp. 308-318).

[6] 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(1), 119.

[7] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Roselander, J. (2019). Towards federated learning at scale: System design. Proceedings of machine learning and systems, 1, 374-388.

[8] Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017, May). Membership inference attacks against machine learning models. In 2017 IEEE symposium on security and privacy (SP) (pp. 3-18).