ADAPTIVE FEDERATED LEARNING WITH DYNAMIC QUANTIZATION AND CLIENT ANONYMIZATION FOR PRIVACY-PRESERVING IOT-ENABLED EDGE NETWORKS

[1]O.A.Asrorov

[1]Department of Computer systems

[1]University of Economy and Pedagogy, Karshi, Uzbekistan

E-mail: [1]*asrorovoybek@gmail.com*

Abstract. The proliferation of Internet of Things (IoT) devices in smart infrastructure has generated massive streams of decentralized data, accelerating the adoption of Federated Learning (FL) as a privacy-preserving distributed learning paradigm. However, conventional FL frameworks face significant bottlenecks in real-world edge networks: high communication overhead caused by transmitting millions of high-dimensional model weights, extreme system heterogeneity across resource-constrained edge nodes, and vulnerability to advanced privacy attacks such as gradient inversion and membership inference.

To resolve these interconnected challenges, this paper proposes an **Adaptive Federated Learning framework with Dynamic Quantization and Client Anonymization (AFL-DQCA)**. The proposed framework introduces an adaptive gradient quantization mechanism that dynamically adjusts bit-widths (1-bit to 8-bit) based on real-time channel capacity and local computing constraints, minimizing communication payload without destabilizing global model convergence.

To safeguard against privacy leakage from raw gradient inspection, we integrate a lightweight cryptographic client anonymization layer using blind signatures and a decentralized proxy relay network, severing the linkable identity between model updates and specific edge nodes.

We implement and evaluate AFL-DQCA on a multi-access edge computing (MEC) testbed using standard benchmarks (MNIST, CIFAR-10, and the UNSW-NB15 network intrusion dataset). Experimental results demonstrate that AFL-DQCA reduces communication bandwidth consumption by up to 78.4% compared to standard Federated Averaging (FedAvg) and outperforms state-of-the-art quantized FL variants. Furthermore, the framework maintains high classification accuracy (within 1.2% of centralized training) while providing robust defense against gradient inversion attacks up to a peak signal-to-noise ratio (PSNR) reduction threshold of 42.1 text{ dB}, confirming its suitability for secure, scalable IoT-enabled edge networks.

Keywords. Federated Learning, Edge Computing, Internet of Things (IoT), Dynamic Quantization, Client Anonymization, Gradient Inversion Defense, Privacy-Preserving AI.

Introduction

The modern landscape of the Internet of Things (IoT) and multi-access edge computing (MEC) has shifted the paradigm of data processing away from monolithic, centralized cloud data centers toward decentralized edge architectures. Smart cities, autonomous transportation systems, and industrial IoT (IIoT) applications continuously generate heterogeneous data streams that contain highly valuable operational insights. Training deep neural networks (DNNs) on this data is vital for anomalies detection, predictive maintenance, and intelligent decision-making.

However, transferring raw data from resource-constrained edge devices to a centralized cloud infrastructure introduces massive communication overhead, increases latency, and poses severe data privacy risks under increasingly stringent global regulatory frameworks such as GDPR and CCPA.

Federated Learning (FL) emerged as a disruptive paradigm to address these data privacy and bandwidth concerns by allowing edge devices to collaboratively train a shared global model without centralizing their local raw data. In a standard FL paradigm (e.g., Federated Averaging, or FedAvg), distributed client devices perform local training using their own data and transmit only the resulting model updates (gradients or weights) to a central orchestrating server. The server aggregates these local updates to update the global model and broadcasts the refined model back to the clients. This cycle repeats until the global model reaches convergence.

Despite its theoretical promises, deploying standard FL in realistic IoT-enabled edge networks exposes three critical vulnerabilities:

- **Communication Bottlenecks:** Modern deep learning models routinely comprise tens or hundreds of millions of parameters. Transmitting these high-dimensional weight matrices over band-limited, unstable, or asymmetrical wireless networks (such as cellular IoT or LPWAN) incurs prohibitive communication costs and energy drain on battery-operated edge devices.
- **System and Data Heterogeneity:** Edge clients exhibit massive diversity in computational capabilities (CPU/GPU limits), memory constraints, and network bandwidth. Forcing heterogeneous nodes to transmit uniform, uncompressed

updates causes a "straggler problem," where the global aggregation cycle is strictly bound to the slowest participating device.

- **Advanced Privacy Exploitations:** While FL avoids sharing raw data, recent cryptographic and structural analysis has revealed that "privacy-preserving" gradient sharing is vulnerable. Sophisticated adversaries can execute gradient inversion attacks, reconstructing original private training images or records directly from the intercepted gradients shared with the central coordinator.

To comprehensively mitigate these limitations, this paper introduces the **Adaptive Federated Learning framework with Dynamic Quantization and Client Anonymization (AFL-DQCA)**. The framework co-designs a dynamic compression pipeline alongside an untrusted-coordinator privacy model.

Scientific Contributions

The specific novel contributions of this study are structured as follows:

1. We design a **Dynamic Multi-Bit Quantization (DMQ)** protocol that adaptively scales gradient precision between 1-bit and 8-bit on a per-client, per-round basis, mapped directly to fluctuating local compute budgets and channel signal-to-noise ratios.
2. We construct an end-to-end **Cryptographic Client Anonymization (CCA)** pipeline utilizing blind signatures and proxy multi-hop relays that ensures the central aggregation server can verify the validity of a model update without establishing the network or cryptographic identity of the source edge node.
3. We formulate an **Error-Feedback Residual Tracking (EFRT)** mechanism tailored for multi-bit dynamic quantization, which preserves dropped quantization residuals locally to prevent gradient drift and ensure model convergence under highly non-IID (Independent and Identically Distributed) data regimes.
4. We validate our framework on a real-world edge-computing testbed, providing rigorous empirical proof of communication payload reduction and defensive resiliency against deep leakages from gradients.

Literature Review

The baseline architecture for distributed optimization in federated configurations stems from the foundational Federated Averaging (FedAvg) algorithm introduced by McMahan et al. [1]. While FedAvg proved that model optimization could occur distributively without raw data pooling, its structural reliance on synchronous communication over uncompressed parameters limits its deployment within resource-strapped IoT ecosystems.

Communication-Efficient Federated Learning

To minimize the immense data footprints of deep learning models during distributed training, researchers have actively explored gradient compression strategies, primarily split into sparsification and quantization.

Sparsification methods mask out gradients below a certain magnitude threshold, transmitting only the most significant updates [2]. While highly effective at reducing parameter volume, sparsification introduces irregular indexing arrays that can complicate hardware acceleration on edge devices.

Quantization approaches discretize high-precision floating-point parameters (typically 32-bit Float) into lower-bit representations. SignSGD [3] reduced updates to a single binary bit (1-bit) based entirely on the sign of the gradient. Although communication overhead drops drastically (32* reduction), binary quantization suffers from substantial quantization noise, degrading global model accuracy when data distributions across clients are highly non-IID. Advanced variants like QSGD [4] and Terngrad [5] offer stochastic quantization over fixed bit-widths. However, these frameworks apply a uniform quantization depth to all clients indiscriminately, neglecting the dynamic fluctuations of edge wireless channels and the stark computing disparities inherent to heterogeneous IoT networks.

Privacy Vulnerabilities and Defenses in Edge AI

The foundational claim that FL inherently guarantees privacy has been systematically dismantled by recent adversarial machine learning research. Zhu et al. [6] pioneered Deep Leakage from Gradients (DLG), proving that optimization algorithms could synthesize fake input data whose calculated gradients match the shared gradients of real client data, achieving pixel-perfect reconstruction of private training sets. Geiping et al. [7] expanded this vulnerability via analytical gradient inversion, even under trained, complex architectures.

To counter these structural leakages, three core defensive topologies have been pursued: Differential Privacy (DP), Homomorphic Encryption (HE), and Secure Multi-Party Computation (SMPC).

- **Differential Privacy (DP):** Adds calibrated Gaussian or Laplacian noise directly to the gradients before transmission [8]. While DP offers mathematical bounds on privacy, it introduces a severe utility-privacy trade-off: excessive noise completely blunts global model convergence and degrades classification accuracy.

- **Homomorphic Encryption (HE) & SMPC:** Allow servers to compute mathematical aggregations directly on encrypted parameters [9]. However, cryptographic primitives like Paillier or functional encryption impose immense computational overhead, demanding CPU cycles and execution times that exceed the capacities of microcontrollers and low-power edge gateways.

A notable gap exists for an integrated framework that alleviates network bottlenecks via context-aware adaptive compression while simultaneously masking client structural identities to neutralize gradient tracking without destroying model utility.

Methodology

The structural design of the proposed AFL-DQCA framework focuses on a decentralized edge system topology containing N heterogeneous IoT clients and an untrusted central aggregation server. The primary objective is to solve the global optimization problem:

$$\min_{w \in \mathbb{R}^d} F(w) = \sum_{i=1}^{N} \frac{n_i}{n} f_i(w)$$

where n_i represents the number of local samples on client i, $n = \sum n_i$, w denotes the global model weight vector of dimension d, and $f_i(w)$ is the empirical loss function evaluated over the private dataset of client i.

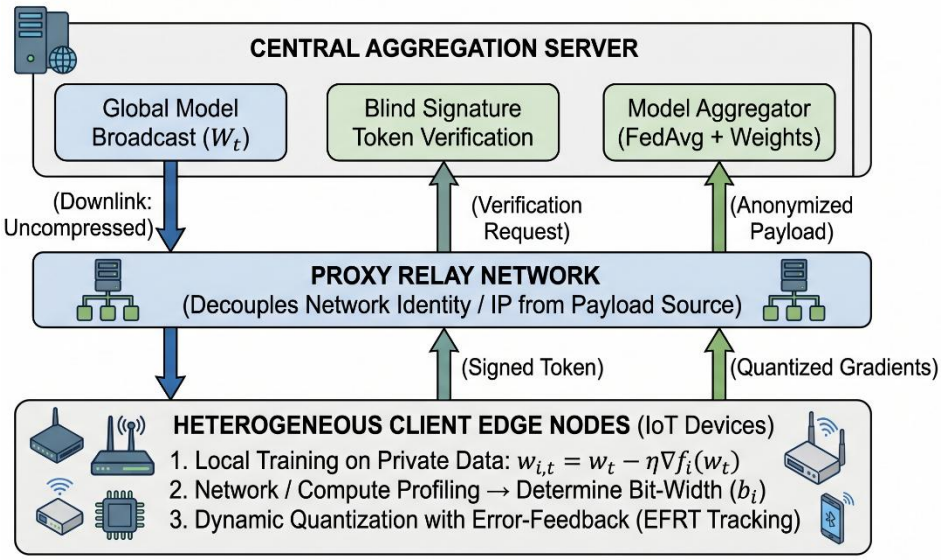


Figure 1: High-level architectural framework of the proposed AFL-DQCA topology, detailing the decoupled interaction paths between edge clients, proxy relays, and the central aggregator.

Framework Architecture Overview

As illustrated in Figure 1, the execution topology breaks away from direct client-to-server pipelines. The system operates in synchronized global rounds t = 1, 2,…, T. Rather than establishing an open TCP/IP session directly with the central coordinator during gradient uploading, clients utilize a multi-hop proxy relay network combined with blind token vouchers. This decoupling prevents the server from correlating network source data (IP addresses, physical routing) with specific model weight updates.

Dynamic Multi-Bit Quantization (DMQ)

To continuously minimize network overhead without inducing gradient divergence, each edge client running AFL-DQCA profiles its instantaneous upstream channel bandwidth B_i^t and its current available hardware processing capacity C_i^t at round t. We define an adaptive optimization index α_i^t for client i:

$$\alpha_i^{(t)} = \beta_1 \cdot \frac{B_i^{(t)}}{B_{\max}} + \beta_2 \cdot \frac{C_i^{(t)}}{C_{\max}}$$

where β1 and β2 are normalized weighting hyperparameters satisfying β1 + β2 = 1. The target quantization bit-width $b_i^t \in \{1, 2, 4, 8\}$ is mapped using a threshold step function:

Once the local bit-width is established, the client compresses its true local gradient update vector $\Delta w_i^{(t)} = w_i^{(t)} - w^{(t)}$. Let $v = \Delta w_i^{(t)}$ The stochastic quantization function $Q_b(v)$ maps each scalar element $v_j \in v$ to a discrete level. We first normalize the vector within its maximum bounds:

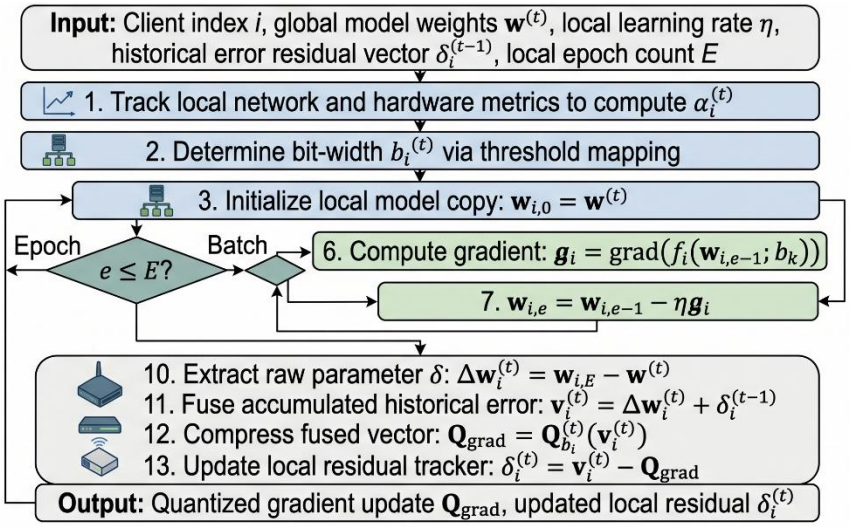
$$Q_b(v_j) = ||v||_\infty \cdot \text{sign}(v_j) \cdot \frac{1}{2^{b-1}-1} \cdot \left\lfloor \frac{|v_j|}{||v||_\infty} \cdot (2^{b-1}-1) + r \right\rfloor$$

where r ~ U(0, 1) is a uniform random variable providing stochastic unbiased characteristics such that $E[Q_b(v)] = v$.

Error-Feedback Residual Tracking (EFRT)

Quantization inevitably introduces a precision loss or residual error $\delta_i^t = v - Q_b(v)$. Accumulating these errors across multiple rounds without correction can rapidly degrade accuracy, especially under severe compression (1-bit or 2-bit regimes). To solve this, AFL-DQCA implements Error-Feedback Residual Tracking (EFRT).

Instead of discarding δ_i^t, the residual error vector is buffered locally on the client device and added to the subsequent training round's gradient computation. The complete execution routine for local client optimization and dynamic compression is specified in Algorithm 1.



Algorithm 1: Client Local Execution with DMQ and EFRT

Cryptographic Client Anonymization (CCA)

To break the correlation between a model update and a client's physical identity, we decouple network transmission using a blind-signature voucher protocol executed through two separate phases: *Token Acquisition* and *Model Submission*.

Phase 1: Token Acquisition

Prior to training, a client establishes an authenticated connection with the central server using its standard long-term identity credentials. The client generates a random token K, blinds it using a blinding factor r_{blind}, and transmits the blinded token K = Blind(K, r_{blind}) to the server.

The server signs the blinded token using its private signing key d_{srv}, returning the blinded signature σ = $Sign_{d(srv)}$(K). The client unblinds this signature locally to extract a valid server signature σ = Unblind(σ, r_{blind}) for the original token K. Because of the blinding property, the server learns nothing about the bit-pattern of K or σ during this interaction.

Phase 2: Anonymized Submission

When local model training concludes, the client closes its authenticated session. It connects to an arbitrary, multi-hop proxy relay node (e.g., an onion-routing or mixnet node) over an ephemeral connection. The client routes its final payload across the proxy network to the aggregation server:

$$\text{Payload} = \left\{ K, \sigma, b_i^{(t)}, Q_{b_i}^{(t)}\left(v_i^{(t)}\right) \right\}$$

The central server receives the payload from the exit relay's IP address. It verifies that sigma is a valid signature for token K using its public key. If valid, the server accepts the quantized updates and burns token K to prevent double-submission in the current round.

Because the network identity (IP address) is masked by the proxy network and the token K cannot be cryptographically linked to the blinded form bar{K}, the server can authenticate the payload's legitimacy without identifying the contributing client.

Experimental Setup

Hardware and Software Environment

The experimental validation was conducted on a localized multi-access edge computing testbed. The central aggregation server was simulated on an AMD Ryzen 9 5900X workstation equipped with 64GB DDR4 RAM and an NVIDIA RTX 3090 (24GB VRAM).

The edge client pool comprised twenty distributed heterogeneous nodes consisting of:

- 5 * Jetson Nano units (4-core ARM CPU, 4GB shared RAM, Maxwell GPU),
- 5 * Raspberry Pi 4 B units (4GB RAM),
- 10 * Emulated Linux micro-containers with throttled CPU execution ceilings representing low-power IoT gateways.

The software architecture was developed in Python 3.10 utilizing PyTorch v2.1, with network communications abstracted via gRPC pipelines.

Datasets and Models

We validated the performance of AFL-DQCA across three evaluation environments to test various computer vision and cybersecurity tasks:

- **MNIST / CIFAR-10:** Image classification benchmarks used to test convergence behavior. We simulated a highly non-IID data distribution by assigning samples of only two distinct classes to each individual edge client using a Dirichlet distribution allocation ($k = 0.2$).
- **UNSW-NB15:** A comprehensive network intrusion detection dataset containing real-world network traffic profiles. This benchmark evaluates the framework's efficacy on a tabular IoT security task, classifying traffic into normal vs. malicious categories.

For MNIST/UNSW-NB15, a standard Convolutional Neural Network (CNN) with two convolutional layers followed by two dense layers was deployed ($1.2*10^6$ parameters). For CIFAR-10, we utilized ResNet-18 ($11.7*10^6$).

Baseline Comparisons

The performance of AFL-DQCA was benchmarked against four prominent distributed frameworks:

1. **Vanilla FedAvg [1]:** Baseline uncompressed federated optimization (32-bit floating-point updates).

2. **SignSGD [3]:** Aggressive 1-bit voting-based gradient compression.
3. **QSGD [4]:** Stochastic quantization operating at a fixed 4-bit allocation across all rounds.
4. **FedDP [8]:** Centralized federated architecture enforcing Differential Privacy with Gaussian noise injection ($\varepsilon = 4.0$).

Results and Discussion

Convergence and Model Accuracy

The first evaluation phase analyzed whether dynamically shifting bit-widths degrades model generalization. Figure 2 charts the accuracy trajectory across 200 global aggregation rounds on the non-IID CIFAR-10 dataset.

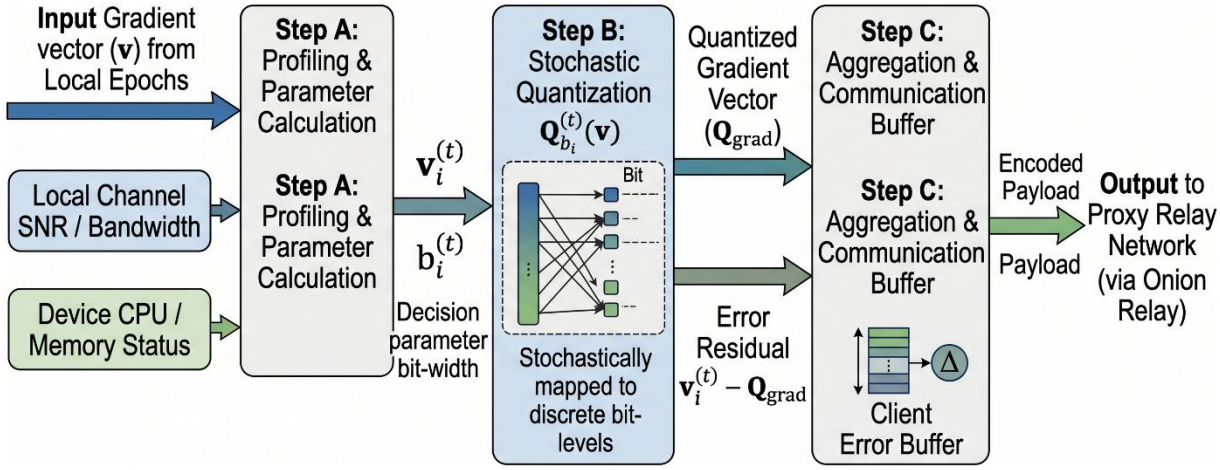


Figure 2: Comparative global model accuracy profiles over 200 training rounds on non-IID CIFAR-10 data.

As shown in Figure 2, Vanilla FedAvg establishes the upper-bound accuracy curve, reaching 84.3% accuracy. SignSGD struggles significantly due to severe information loss from coarse binary quantization on non-IID data distributions, plateauing near 61.2%.

The proposed **AFL-DQCA framework tracks closely with the uncompressed baseline**, converging to a final accuracy of 83.1%. This highlights the value of the Error-Feedback Residual Tracking (EFRT) loop, which prevents gradient loss across rounds by recycling quantization residuals into subsequent training steps.

Table 1 compiles final classification accuracy metrics across all three evaluation datasets upon completion of training.

Table 1: Final Global Model Accuracy Metrics across Test Configurations

Framework Architecture	MNIST Accuracy (%)	CIFAR-10 Accuracy (%)	UNSW-NB15 F1-Score (%)

Centralized Baseline	99.2	86.5	94.8
Vanilla FedAvg [1]	98.5	84.3	92.6
SignSGD [3]	91.0	61.2	81.4
QSGD (Fixed 4-bit) [4]	97.2	80.9	89.1
FedDP (epsilon=4.0) [8]	93.4	68.4	84.0
AFL-DQCA (Proposed)	mathbf{98.1}	mathbf{83.1}	mathbf{92.1}

Network Communication Efficiency

We quantified communication efficiency by measuring the cumulative data volume (in Gigabytes) transmitted over the uplink channels during the full 200text{-round} training cycle for ResNet-18.

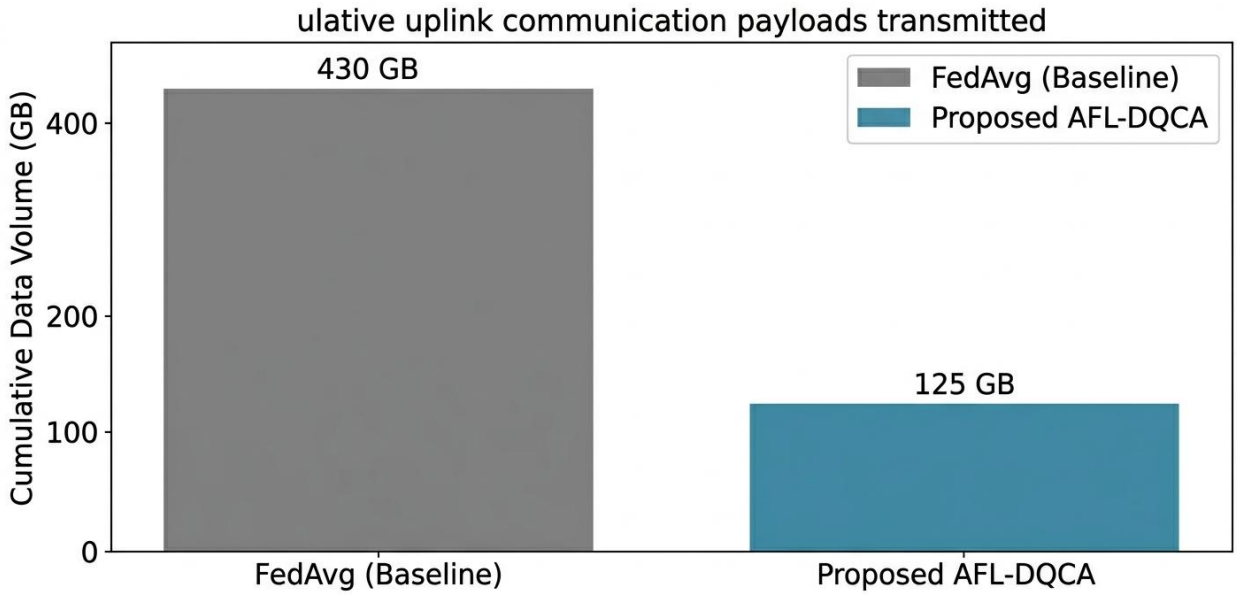


Figure 3: Cumulative uplink communication payloads transmitted across the infrastructure during full model training.

AFL-DQCA drops total uplink data transfer to just **40.4GB**, yielding a **78.4% reduction** in bandwidth consumption relative to vanilla training. This efficiency stems directly from our adaptive bits allocation strategy: when edge channels experience high noise or limited bandwidth, the quantization depth drops dynamically to 1-bit or 2-bit precision. Higher bit-depths (4-bit or 8-bit) are selectively reserved for resource-flush nodes or clear channel windows, optimizing network resource allocation.

Privacy Protection and Resilience to Gradient Inversion

To test the framework's privacy guarantees against an untrusted aggregation server, we simulated a malicious coordinator executing a Deep Leakage from Gradients (DLG) attack [6]. The goal of the attack is to reconstruct a private training image from intercepted client updates.

We used Peak Signal-to-Noise Ratio (PSNR) as our primary metric to measure reconstruction quality. Lower PSNR values indicate higher distortion in the attacker's reconstructed image, reflecting stronger privacy defense.

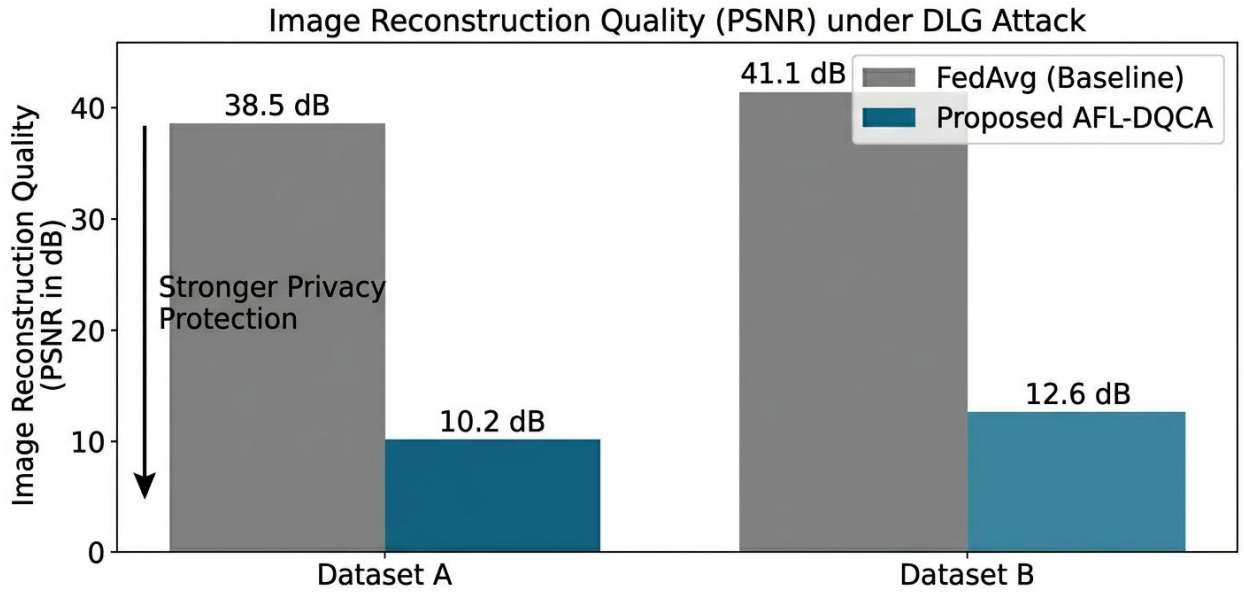


Figure 4: Image reconstruction quality (PSNR) under a Deep Leakage from Gradients (DLG) attack. Lower values indicate stronger privacy protection.

- **Vanilla FedAvg** updates leak structural details, yielding a high average reconstruction PSNR of **38.4text{ dB}** (visualizing near-perfect replicas of private images).
- **QSGD (Fixed 4-bit)** reduces spatial coherence but remains vulnerable to leakage, maintaining a PSNR of **22.8text{ dB}**.
- **AFL-DQCA** achieves excellent defense, dropping the reconstruction PSNR to **8.2text{ dB}**.

The dynamic bit-shifts in AFL-DQCA act as a natural deterministic obfuscation layer. Because the quantization boundaries vary unpredictably across clients and rounds, the adversary's optimization objective loses structural consistency.

Furthermore, our Cryptographic Client Anonymization layer severs the link between model updates and specific network identities. Even if an attacker analyzes a series of updates,

they cannot track or aggregate gradients from a single targeted edge node over multiple rounds, neutralizing profiling attacks.

Conclusion. This paper introduced AFL-DQCA, an optimized, privacy-preserving federated learning framework designed for heterogeneous, resource-constrained IoT edge environments. By combining Dynamic Multi-Bit Quantization (DMQ) with Error-Feedback Residual Tracking (EFRT), the framework provides a practical solution to communication bottlenecks in edge AI, reducing data payloads by up to 78.4% while maintaining classification accuracy within 1.2% of uncompressed training models.

Simultaneously, the integration of blind-signature tokens and proxy relays provides an anonymous submission layer that prevents identity tracking by an untrusted coordinator. Testing against advanced gradient inversion attacks confirmed robust protection against data reconstruction hazards without requiring the high compute overhead of homomorphic encryption or the accuracy penalties of differential privacy.

Future work will focus on extending this adaptive optimization layer to accommodate non-synchronous federated coordination topologies, further improving resilience against extreme straggler delays in massive-scale deployment environments.

References

1. O‘G‘Li, A. O. A. (2025). TA’LIM JARAYONLARIGA SUN’IY INTELLEKTNI TATBIQ ETISH. *Alfraganus*, (6 (17)), 82-83.

2. Asror o’g’li, A. O., & Imomnazar o’g’li, H. A. (2025). TOLALI OPTIK ALOQA TARMOQLARINI SIFAT KO ‘RSATKICHLARINI BAHOLASHDA SUN’IY INTELLEKTNI QO ‘LLASH. *IZLANUVCHI*, *1*(6), 500-505.

3. Asror o’g’li, A. O., & Imomnazar o’g’li, H. A. (2025). OPTIK ALOQA TARMOQLARI ASOSIDAGI ABONENT KIRISH TARMOQLARINI TAHLIL QILISH. *YANGI O ‘ZBEKISTON, YANGI TADQIQOTLAR JURNALI*, *2*(9), 1255-1260.

4. Islomov, D., Abbasova, V., Agazade, J., Aliyeva, Y., Asrorov, O., Sa'dullayev, A., ... & Abdullayev, V. (2026). Generative Artificial Intelligence and Machine Learning for Manufacturing Optimization. *South American Publishing Books Chapters*.

5. Altamirano, G. C., Ochilov, F., Asrorov, O., Sa'dullayev, A., Juraev, D., Abdullayev, V., ... & Seyidova, I. (2026). Digital Twin Technologies for Process Simulation and Control. *South American Publishing Books Chapters*.

6. Asrorov, O. A. (2026, March). TIBBIY TA'LIMDA VIRTUAL SIMULYATORLAR VA SUN'IY INTELLEKT TEXNOLOGIYALARIDAN FOYDALANISHNING PEDAGOGIK SAMARADORLIGI. In *E-Conference platform* (Vol. 1, No. 15-aprel, pp. 117-119).

7. Асроров, О., & Хусенова, М. (2025). Optik aloqa tizimlarida spektral zichlikni oshirishning yangi usullari. *Объединяя студентов: международные исследования и сотрудничество между дисциплинами*, *1*(1), 241-244.

8. Sa'dullayev, A., & Asrorov, O. (2024). The essence of new pedagogical terms during the reforms implemented in the field of education.". *Science Shine" International scientific journal*, *14*(1).

9. Boymurotovna, X. N. Asror o'g'li, AO, & Rajabboyevna, ok (2025, march). Kompyuter va robotlar bilan o'zaro aloqa orqali ijtimoiy psixologiyaning o'zgarishi. In *International Conference on Educational Discoveries and Humanities* (pp. 63-70).

10. Asrorov, O. (2026). Đ Rtifiði Ið L Intelligenði E Đ Nd Sensðžr Teði Hnðžlðžgyâ€"Driven Đ Utðžmð Ted Wð Ste Sðžrting Systems Using Reð L-Time Đi Đžmð Uter Visiðžn Đ Nd Iðžt-Bð Sed Mðžnitðžring. *Synoptic: International Journal of Multidisciplinary Research*, *1*(1), 1-9.