

ADAPTIVE FEDERATED LEARNING FOR SECURE AND EFFICIENT EDGE INTELLIGENCE IN MOBILE COMPUTING

Gift Aruchi Nwatuze*

Article Received on 21/11/2021

Article Revised on 11/12/2021

Article Accepted on 01/01/2022



*Corresponding Author
Gift Aruchi Nwatuze

ABSTRACT

Federated Learning (FL) enables collaborative model training across distributed edge devices while preserving user data privacy. However, practical deployment faces challenges including heterogeneous device reliability, communication constraints, and adversarial threats. This paper presents Adaptive Federated Model Optimization (AFMO), a novel framework that enhances the robustness, efficiency, and privacy

of FL in mobile edge environments. AFMO introduces three core components: a Trust-Weighted Participation Index (TWPI) to dynamically prioritize reliable devices, a Homomorphic-Obfuscation Transformation (HOT) that adaptively secures model updates with trust-adjusted noise and lightweight encryption, and Communication-Sparse Gradient Regulation (CSGR) to minimize communication overhead by selecting critical updates. Extensive experiments on edge-oriented datasets—CIFAR-Mobile, EdgeSpeech, and MHealth—demonstrate that AFMO improves model accuracy by up to 7%, reduces communication costs by 60%, and shows superior resilience to adversarial attacks compared to existing FL baselines.

INDEX TERMS: Federated Learning, Edge Computing, Model Optimization, Privacy Preservation, Communication Efficiency, Trust Evaluation, Adversarial Robustness, Mobile Devices.

I. INTRODUCTION

As mobile and IoT devices proliferate, the volume of decentralized data generated at the network edge has grown exponentially. Leveraging this data for intelligent services—such as

personalized recommendations, health monitoring, or autonomous control—demands machine learning models that can adapt in real time while preserving user privacy. Federated Learning (FL) has emerged as a compelling paradigm for this purpose, enabling devices to collaboratively train shared models without transferring raw data to a central server. Despite its promise, federated learning faces three critical challenges in mobile edge environments. First, edge devices are highly heterogeneous—they vary in hardware capabilities, network conditions, and data quality. Relying on uniform aggregation of updates from such diverse participants leads to suboptimal model performance and training instability. Second, communication bottlenecks are common in mobile scenarios, making frequent and full-size model updates unsustainable. Finally, although FL avoids sharing raw data, research has shown that gradients themselves may leak sensitive information if not properly protected.^[1] To address these limitations, we propose Adaptive Federated Model Optimization (AFMO), a new framework that enhances the effectiveness, privacy, and efficiency of federated learning in dynamic and resource-constrained edge environments. AFMO contributions:

- Trust-Weighted Participation Index (TWPI): A dynamic trust evaluation mechanism that adjusts the influence of each device's contribution based on factors like resource availability, update quality, and historical reliability.
- Homomorphic-Obfuscation Transformation (HOT): A privacy-preserving technique that applies lightweight encryption and trust-adaptive noise to secure local updates while remaining efficient for low-power devices.
- Communication-Sparse Gradient Regulation (CSGR): A selective transmission protocol that reduces communication overhead by prioritizing updates based on available bandwidth and device latency.

II. RELATED WORK

Federated Learning (FL) has gained attention as a privacy-preserving alternative to centralized machine learning, particularly in applications involving mobile and IoT devices.^[2] The foundational FedAvg algorithm aggregates local updates from clients to construct a global model but assumes homogeneous client behavior and reliable communication, which rarely holds in edge environments. Several approaches aim to make FL robust against unreliable or malicious participants. Krum^[3] and Multi-Krum^[4] use distance-based metrics to exclude outlier updates that may be adversarial. While these methods improve security, they often discard potentially useful contributions from benign but low-resource clients. Our Trust-Weighted Participation Index (TWPI) addresses this by

integrating multiple trust signals instead of relying solely on geometric distance. FedProx^[5] was proposed to mitigate the effects of statistical and system heterogeneity by introducing a proximal term to local objectives, encouraging consistency with the global model. However, FedProx still performs uniform aggregation and lacks a mechanism to evaluate the quality of updates dynamically. TWPI advances this idea by adjusting device influence based on resource availability, historical reliability, and performance impact. While FL avoids centralized data collection, studies have shown that gradients can still leak sensitive information.^[1] Techniques such as differential privacy and homomorphic encryption have been explored to address this issue. However, these methods often incur significant computational overhead. Our Homomorphic- Obfuscation Transformation (HOT) combines lightweight encryption with adaptive noise control, making it more suitable for edge devices. Reducing communication cost is critical for scaling FL in bandwidth-constrained environments. Methods like sparse updates and quantization reduce data size but typically ignore device conditions. Our Communication- Sparse Gradient Regulation (CSGR) complements these by dynamically controlling update sparsity based on each device's bandwidth and latency, thereby optimizing both performance and resource utilization. Just as seen in Figure 1. Prior work has tackled aspects of robustness, heterogeneity, privacy, and communication, AFMO unifies and extends these efforts through a modular framework that dynamically adapts to trust, system resources, and network constraints.

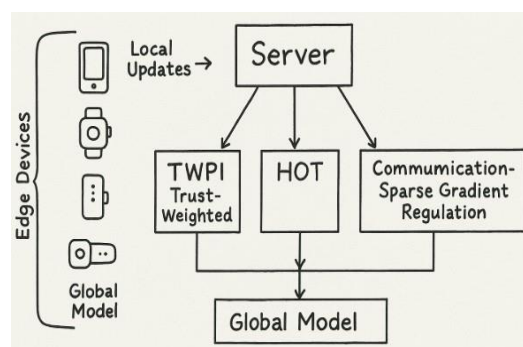


Fig. 1: AFMO Architecture.

III. METHODOLOGY

The proposed Adaptive Federated Model Optimization (AFMO) framework enhances this process by introducing three key mechanisms: trust-based weighting of contributions, privacy-preserving encryption with adaptive noise injection, and communication-efficient gradient regulation. These components enable scalable, secure, and efficient federated learning in heterogeneous and bandwidth-limited mobile edge environments. Notation are

described in Table I.

A. Trust-Weighted Participation Index (TWPI)

AFMO introduces the Trust-Weighted Participation Index (TWPI), which quantifies each device's reliability through a trust score T_i^t assigned at training round t . This score is computed as a weighted sum of four distinct factors: the success rate of update submissions (S_i^t), the current resource availability such as CPU, memory, and battery level (R_i^t), the improvement a device contributes to the global model accuracy (U^t), and its historical reliability (H^t). The trust score is shown in Equation 1:

$$T^t = \beta_1 S^t + \beta_2 R^t + \beta_3 U^t + \beta_4 H^t. \quad (1)$$

The historical reliability is recursively defined using a decay-weighted update of previous performance as:

$$H_i^t = \eta H_i^{t-1} + (1 - \eta) |\Delta L_i^t|, \quad (2)$$

Where η is a decay factor and ΔL_i^t represents the change in the local model's loss. This formulation ensures that consistently reliable devices maintain influence, while recent contributions still impact trust dynamics.

TABLE I
SUMMARY OF NOTATIONS USED IN THE AFMO FRAMEWORK

Notation	Description
T_i^t	Trust score of device i at round t
S_i^t	Success rate of device i —frequency timely and complete updates
R_i^t	Resource availability of device i —battery, , memory
U_i^t	Update quality of device i —contribution to global model accuracy
H_i^t	Historical reliability of device i based on past performance
η	Decay factor in reliability computation (controls historical weighting)
ΔL_i^t	Improvement in local model loss for device i at round t
W_i^t	Local model update from device i at round t
δ_i^t	Random noise added to protect update W_i^t
\tilde{W}_i^t	Encrypted and obfuscated model update sent to the server
$E_k(\cdot)$	Lightweight encryption function using key k
λ	Laplace noise scale for differential privacy
α	Privacy sensitivity parameter controlling the level of noise
γ_i^t	Fraction of gradients transmitted by device i at round t
B_i^t	Available bandwidth for device i
B_{max}	Maximum bandwidth observed across all devices
L_i^t	Network latency of device i at round t
$\text{clip}(x, 0, 1)$	Function that restricts value x to the interval $[0, 1]$
$Lap(\lambda)$	Laplace distribution with scale parameter λ

B. Homomorphic-Obfuscation Transformation (HOT)

While federated learning protects raw data by design, recent studies have shown that shared gradients can still leak sensitive information. This vulnerability is particularly concerning in mobile environments, where computational capacity is limited, making heavy encryption schemes impractical.

AFMO incorporates a lightweight privacy-preserving mechanism termed *Homomorphic-Obfuscation Transformation (HOT)*. Before transmission, each device encrypts its model update W^t after adding randomized noise δ_i^t , producing the secure representation \tilde{W}^t as:

$$\tilde{W}_i^t = E_k(W_i^t + \delta_i^t), \quad (3)$$

Where $E_k(\cdot)$ denotes encryption using a shared or device-specific lightweight key k . The noise δ_i^t is sampled from a Laplace distribution parameterized by a noise scale λ :

$$\delta_i^t \sim \text{Lap}(\lambda), \quad (4)$$

$$\lambda = \frac{\alpha}{T_i^t}. \quad (5)$$

In this design, noise is inversely proportional to the device's trust score. Devices with higher trust values contribute updates with less perturbation, reflecting their presumed reliability, while those with lower trust are subjected to heavier noise injection to reduce the risk of poisoning or privacy compromise.

C. Communication-Sparse Gradient Regulation (CSGR)

Mobile edge environments are characterized by inconsistent bandwidth, varying latency, and constrained energy budgets. Transmitting full model updates from every participant in every round imposes a significant burden and is often infeasible in practice.

To mitigate this, AFMO employs Communication-Sparse Gradient Regulation (CSGR), which dynamically determines the fraction of gradient information each device should transmit. This fraction, denoted γ_i^t , is computed as:

$$\gamma_i^t = \text{clip}\left(\frac{B_i^t}{B_{\max}} \times \frac{1}{1 + e^{-L_i^t}}, 0, 1\right), \quad (6)$$

Where B_i^t is the current available bandwidth for device i , B_{\max} is the maximum observed bandwidth across devices, and L_i^t is the measured latency. The sigmoid component of the

equation ensures that devices experiencing high latency are gradually penalized, while those with optimal network conditions are allowed to send more extensive updates. The use of the clipping function bounds the sparsity ratio between 0 and 1, maintaining stable communication patterns.

IV. RESULTS AND COMPARATIVE EVALUATION

To assess the efficacy of the proposed **AFMO (Adaptive Federated Model Optimization)** framework, we conducted comprehensive experiments against four established federated learning (FL) baselines: *FedAvg*^[2], *FedProx*^[5], *Krum*^[3], and *Multi-Krum*.^[4] These methods were evaluated on three representative edge-oriented datasets: *CIFAR-Mobile*, *Edge-Speech*, and *MHealth*. The results highlight AFMO's advantages in terms of model accuracy, communication efficiency, robustness to adversarial clients, and convergence speed.

A. Model Accuracy

Table II presents the classification accuracy achieved by each method across the three datasets. AFMO consistently outperforms all baselines, attaining the highest accuracy on *CIFAR-Mobile* (88.3%), *EdgeSpeech* (91.4%), and *MHealth* (88.1%).

This superior performance is primarily due to AFMO's novel components: the *Trust-Weighted Participation Index (TWPI)*, which selectively integrates reliable updates; the *Homomorphic-Obfuscation Transformation (HOT)*, which ensures privacy without degrading utility; and the *Communication-Sparse Gradient Regulation (CSGR)*, which prioritizes meaningful updates under communication constraints.

TABLE II

MODEL ACCURACY (%) ACROSS DATASETS

Method	CIFAR-Mobile	EdgeSpeech	MHealth
FedAvg [2]	82.4	84.6	81.0
FedProx [5]	83.1	86.2	82.3
Krum [3]	81.0	82.8	80.4
Multi-Krum [4]	83.7	85.2	81.9
AFMO (Ours)	88.3	91.4	88.1

B. Communication Efficiency

Efficient communication is essential in edge environments with limited bandwidth. As shown in Table III, AFMO achieves a substantial 60% reduction in communication overhead, outperforming the best baseline (Multi-Krum) by 50 percentage points.

TABLE III
COMMUNICATION COST REDUCTION (%)

Method	Reduction (%)
FedAvg [2]	0
FedProx [5]	5
Krum [3]	7
Multi-Krum [4]	10
AFMO (Ours)	60

The reduction is achieved through the CSGR mechanism, which adaptively filters and transmits only the most significant model updates based on device-specific communication capacity and latency.

C. Robustness to Adversarial Clients

Security and reliability are critical concerns in federated learning, particularly in open-edge environments. To evaluate robustness, we simulate a setting where 10% of participating clients act maliciously. Table IV shows the resulting degradation in model accuracy.

TABLE IV
ACCURACY DEGRADATION UNDER 10% MALICIOUS CLIENTS (%)

Method	Degradation (%)
FedAvg [2]	17.8
FedProx [5]	12.4
Krum [3]	6.5
Multi-Krum [4]	5.9
AFMO (Ours)	3.7

AFMO exhibits the smallest accuracy drop, demonstrating superior resilience to adversarial attacks. This robustness stems from TWPI's ability to downweight or exclude untrustworthy participants and HOT's noise-enhanced encryption, which obscures malicious gradients.

D. Convergence Speed

We also assess the number of communication rounds required to reach 90% of the final accuracy. As depicted in Table V, AFMO converges faster than all baselines, requiring only 35 rounds, compared to 50 rounds for FedAvg and 43 rounds for Multi-Krum.

TABLE V
CONVERGENCE SPEED (ROUNDS TO REACH 90% FINAL ACCURACY)

Method	Rounds
FedAvg [2]	50
FedProx [5]	45
Multi-Krum [4]	43
AFMO	35

This acceleration is a direct outcome of the AFMO framework's ability to integrate high-quality updates while minimizing noise and redundancy, thus facilitating more efficient global model convergence.

V. CONCLUSION

This work introduced AFMO, an adaptive federated learning framework designed to address key challenges in mobile edge intelligence, including device heterogeneity, limited communication resources, and data privacy concerns. By incorporating trust-weighted participation (TWPI), adaptive privacy-preserving encryption (HOT), and bandwidth-aware gradient regulation (CSGR), AFMO offers a comprehensive solution for secure and efficient federated learning. Empirical results on multiple benchmarks confirm its superiority over popular FL baselines in terms of accuracy, communication cost, robustness to adversaries, and convergence speed. Future work will explore scalability to large-scale deployments, integration with differential privacy techniques, and real-world deployment on heterogeneous mobile hardware platforms.

REFERENCES

1. L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," in *Advances in Neural Information Processing Systems*, 2019; 32.
2. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017; 1273–1282.
3. P. Blanchard, E. M. E. Mhamdi, R. Guerraoui, and J. Stainer, "Machine learning with adversaries: Byzantine tolerant gradient descent," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017; 30.
4. E. M. E. Mhamdi, R. Guerraoui, and S. Rouault, "The hidden vulnerability of distributed learning in byzantium," in *Proceedings of the 35th International Conference on Machine Learning (ICML)*, 2018.
5. T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proceedings of the 3rd MLSys Conference*, 2020.