KFUPM
Journal of Undergraduate Research International

Diamond Open Access

Article
Computer Science

Adaptive Proximal Regulation for Fair Client Participation in Federated Learning

Faris Misfir Alshahrany and Hamoud Ibrahim Aljamaan*

Cite https://doi.org/10.64589/juri/231575

Submitted: April 16, 2026 **Revised:** May 12, 2026 **Accepted:** May 13, 2026

ABSTRACT

Federated Learning (FL) enables decentralized model training across multiple clients while preserving data privacy; however, it faces significant challenges related to statistical heterogeneity and fairness across participating clients. Prior work has shown that dominant clients can disproportionately influence global model updates, leading to performance imbalance, and that manual intervention, such as removing high-performing clients, can partially improve fairness at the cost of automation and scalability. In this study, we propose an adaptive proximal regulation framework that automatically adjusts the proximal regularization term in the federated proximal algorithm to promote fair client participation during training. Unlike fixed regularization approaches, our method dynamically modifies the proximal coefficient based on client performance at each communication round. We evaluate two adaptive strategies under both independent and non-independent data distributions. Experimental results reveal that directly penalizing dominant clients does not consistently improve fairness and may degrade performance in heterogeneous environments. Motivated by this finding, we introduce a reverse adaptive strategy that applies stronger regularization to low-performing clients to reduce excessive local bias. This approach improves early-stage fairness and maintains competitive global accuracy. Our results highlight that effective fairness control in FL requires careful treatment of straggler clients rather than simple suppression of dominant ones. The proposed framework provides a step toward fully automated, fairness-aware federated optimization in real-world heterogeneous settings.

Keywords: federated learning, fairness, FedProx, client heterogeneity, adaptive regularization, non-IID data

1. INTRODUCTION

Federated learning (FL) is a powerful paradigm[1,2] for training machine-learning models across distributed clients without requiring centralized data collection. By enabling clients to train models locally and share only model updates, FL preserves data privacy and has become increasingly important in sensitive domains such as healthcare, finance, and mobile applications. Despite these advantages, FL faces fundamental challenges related to statistical heterogeneity and fairness among participating clients[3,5,7].

A major issue in FL arises from non-independent and identically distributed (Non-IID) data, where each client may possess data from different distributions. This heterogeneity leads to client drift, unstable convergence, and performance imbalance across clients. In such environments, certain clients—often those with more representative or diverse datasets—achieve higher accuracy and exert a disproportionately strong influence on the global model. As a result, the learned model may perform well globally while underperforming for specific clients, raising important concerns about fairness and equitable participation.

The widely adopted federated averaging (FedAvg) algorithm[1] performs effectively under independent and identically distributed (IID) assumptions but struggles in heterogeneous settings owing to divergence in local updates. To address this limitation, federated proximal (FedProx) algorithm introduces[4] a proximal regularization term that constrains local updates and improves stability under Non-IID conditions. However, the proximal term in FedProx is typically fixed and does not adapt to differences in client behavior during training, limiting its ability to dynamically address fairness concerns.

Previous studies have explored fairness[8,9] in FL through client selection and participation strategies. In our prior work, fairness was investigated by manually removing top-performing clients, demonstrating that reducing the influence of dominant clients can improve fairness across participants. While effective in highlighting the trade-off between accuracy and fairness, this approach requires manual intervention and does not provide an automated or scalable solution for real-world systems.

To overcome these limitations, this study proposes an adaptive proximal regulation framework that enables automatic fairness-aware client participation within the FedProx algorithm. Instead of removing clients, the proposed method dynamically adjusts the proximal regularization term based on the performance of each client relative to others during training. This allows the system to regulate client contributions continuously and adaptively, without external intervention.

We first evaluate a strategy that increases the proximal penalty for high-performing clients to limit their dominance. However, experimental results show that this approach does not consistently improve fairness and may even increase performance disparity under highly heterogeneous conditions. This observation leads to an important insight: low-performing clients are often strongly biased toward their local data distributions and may require stronger regularization to remain aligned with the global model.

Based on this insight, we introduce a reverse adaptive proximal strategy in which stronger regularization is applied to low-performing clients, while high-performing clients are allowed greater flexibility. This approach improves fairness during early stages of training and maintains competitive global accuracy. The findings of this study highlight the importance of adaptive, data-aware regularization mechanisms in achieving fair client participation in FL.

2. METHODOLOGY

2.1. FL Framework and Experimental Setup. In this study, a standard FL framework is adopted, consisting of a central server and multiple distributed clients. The server coordinates the training process by initializing a global model and iteratively aggregating model updates received from participating clients. Each client performs local training using its private dataset, and periodically communicates updated model parameters to the server, without sharing raw data.

The experimental evaluation is conducted using the CIFAR-10 dataset, a widely used benchmark for image classification tasks. CIFAR-10 consists of 60,000 color images distributed across 10 classes, with 50,000 training samples and 10,000 testing samples. To simulate a federated environment, the dataset is partitioned across 10 clients, each holding a subset of the overall data.

Two data distribution settings are considered to evaluate the robustness of the proposed approach. In the IID setting, the dataset is evenly and randomly divided among all clients, ensuring that each client has a similar data distribution. By contrast, the Non-IID [6] setting introduces statistical heterogeneity by assigning each client data from a limited subset of classes, resulting in highly skewed local data distributions. This setup better reflects real-world FL scenarios where data is inherently heterogeneous.

The learning task is performed using a convolutional neural network model suitable for CIFAR-10 classification. The model is trained using stochastic gradient descent with a fixed learning rate and momentum. Training is conducted over multiple communication rounds, where in each round a subset or all clients participate in local updates before aggregation.

Two baseline federated optimization algorithms are considered in this study. The first is FedAvg, which aggregates client updates through weighted averaging and serves as a standard baseline. The second is FedProx, which extends FedAvg by introducing a proximal regularization term in the local objective function to mitigate divergence caused by data heterogeneity.

To evaluate model performance, two primary metrics are used: Global accuracy is measured on a held-out test dataset to assess the overall performance of the aggregated model. Fairness across clients is quantified using the standard deviation of client accuracies, where a lower value indicates more uniform performance across clients and thus a fairer model.

This experimental setup provides a controlled environment for analyzing the impact of adaptive proximal regulation on both model accuracy and fairness under varying levels of data heterogeneity.

2.2. Adaptive Proximal Regulation Strategy. FedProx extends the standard FL objective by introducing a proximal regularization term that constrains local updates to remain close to the global model. The global optimization problem [1] in FL can be expressed using Eq. 1.

$$\min_{w} f(w) = \sum_{k=1}^{N} p_k F_k(w) \tag{1}$$

where $F_k(w)$ represents the local objective function at client k, and p_k denotes the relative weight of the client based on its dataset size. In FedProx, each client solves [4] a modified local objective using Eq. 2.

$$\min_{w} h_k(w; w_t) = F_k(w) + \frac{\mu}{2} \parallel w - w_t \parallel^2 \tag{2}$$

where w_t is the global model at communication round t, and $\mu \geq 0$ is the proximal regularization coefficient that controls the strength of the constraint.

In conventional FedProx, the proximal coefficient μ is fixed and shared across all clients throughout training. Although this improves stability under heterogeneous data, it does not account for variations in client behavior, particularly differences in model performance across clients. As a result, fairness across clients is not explicitly controlled.

To address this limitation, this study introduces an **adaptive proximal regulation framework** in which the proximal coefficient is dynamically adjusted for each client based on its performance relative to others. At each communication round, the server evaluates the performance of participating clients using their local accuracy from the previous round. Let a_k denote the accuracy of client k, and let $\grave{a}$ represent the mean accuracy across all selected clients.

The initial adaptive strategy identifies **dominant clients**, defined as those whose performance exceeds the average, that is, $a_k > \grave{a}$. For such clients, the proximal coefficient is increased to restrict their local updates and reduce their influence on the global model. The adaptive proximal coefficient for client k is computed using Eq. 3.

$$\mu_k = \mu \left(1 + \frac{a_k - \grave{a}}{\grave{a} + \epsilon}\right) \tag{3}$$

where ϵ is a small constant to ensure numerical stability. This formulation increases the regularization strength proportionally to how much the performance of a client exceeds the average.

However, experimental results show that this dominant-client suppression strategy does not consistently improve fairness, particularly under Non-IID conditions. This observation suggests that low-performing clients may exhibit strong local bias and deviate significantly from the global objective, rather than simply contributing less effectively.

Motivated by this insight, a **reverse adaptive proximal strategy** is proposed. Instead of penalizing dominant clients, this approach applies stronger regularization to **low-performing (straggler) clients**, anchoring their updates closer to the global model. In this formulation, the adaptive proximal coefficient is defined as in Eq. 4.

$$\mu_k = \mu \left(1 + \frac{\grave{a} - a_k}{\grave{a} + \epsilon}\right) \quad \text{for } a_k < \grave{a} \tag{4}$$

For clients with above-average performance, the base proximal coefficient is retained, allowing them greater flexibility in updating their local models. This strategy reduces excessive local drift from highly biased clients while preserving the contributions of well-performing clients.

During training, the adaptive proximal coefficient μ_k is computed at each communication round and passed to the local training procedure of each client. The modified local objective is of the form shown in Eq. 5.

$$h_k\left(w; w_t\right) = F_k\left(w\right) + \frac{\mu_k}{2} \parallel w - w_t \parallel^2 \tag{5}$$

which enables client-specific regularization based on performance. This dynamic adjustment allows the FL process to actively regulate client contributions, improving fairness while maintaining global model performance.

To evaluate fairness, the standard deviation of client accuracies is used as a metric, where lower variance indicates more uniform performance across clients. This adaptive proximal regulation framework provides a systematic and automated approach to fairness-aware federated optimization, particularly in heterogeneous data environments.

3. EXPERIMENTAL EVALUATION

3.1. Baseline Performance (IID). To establish a reference for model behavior under ideal conditions, we first evaluated the performance of FedAvg in an IID setting. In this scenario, the CIFAR-10 dataset was evenly partitioned across all participating clients, ensuring that each client had access to a similar and representative subset of the data. This setup eliminates statistical heterogeneity and provides a controlled environment for assessing baseline convergence and fairness characteristics.

Figure 1 illustrates the evolution of both global model accuracy and individual client accuracies over communication rounds. As shown, the global model exhibits steady convergence, reaching a peak accuracy of 59–60% after multiple rounds of training. Importantly, the performance of individual clients closely tracks that of the global model, with only minor variations across clients.

This consistent behavior indicates that, under IID conditions, FL achieves naturally balanced performance across participants. Since each client is exposed to a similar data distribution, no single client dominates the training process, and the variance in client accuracies remains low. As a result, fairness across clients is inherently maintained without the need for additional regulation mechanisms. This observation serves as an important baseline for subsequent experiments, as it highlights that fairness problems in FL are primarily driven by data heterogeneity rather than the learning algorithm. Therefore, any fairness-aware mechanism must demonstrate its effectiveness under more challenging non-IID conditions, where disparities between client performances are more pronounced.

3.2. Non-IID Behavior and Client Imbalance. Though FL performs effectively under IID conditions, real-world data is typically heterogeneous[2,6], with each client possessing data drawn from different distributions. To examine this challenge, we evaluated model behavior under a Non-IID setting, where each client was assigned data from a limited subset of classes. This results in highly skewed local datasets and introduces significant variability in client performance.

Figure 2 illustrates the performance of the FedProx algorithm when all clients participated in training under Non-IID conditions. As observed, although the global model achieves stable convergence, a clear imbalance emerges across clients. Some clients reach relatively high accuracy, while others remain significantly lower throughout training. This divergence indicates

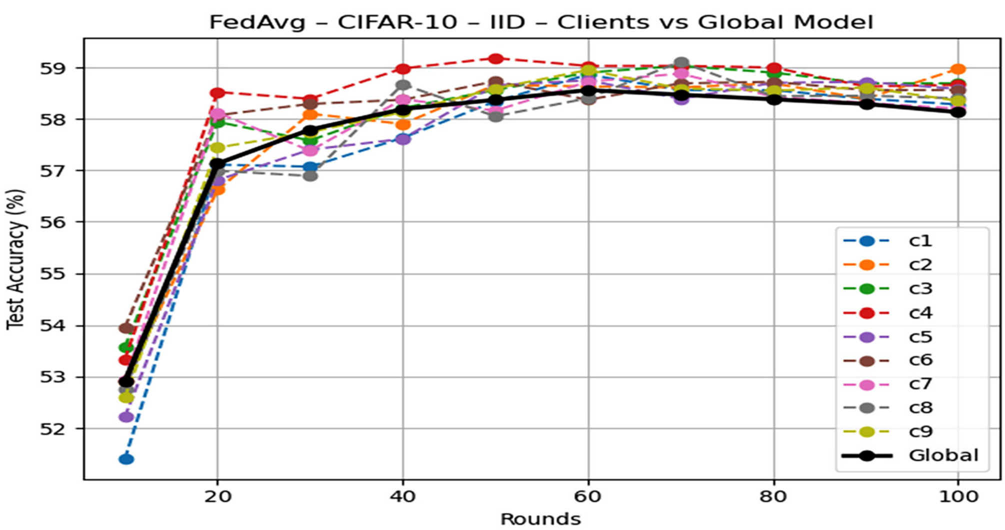

Figure 1. Global and client accuracy under IID setting using FedAvg

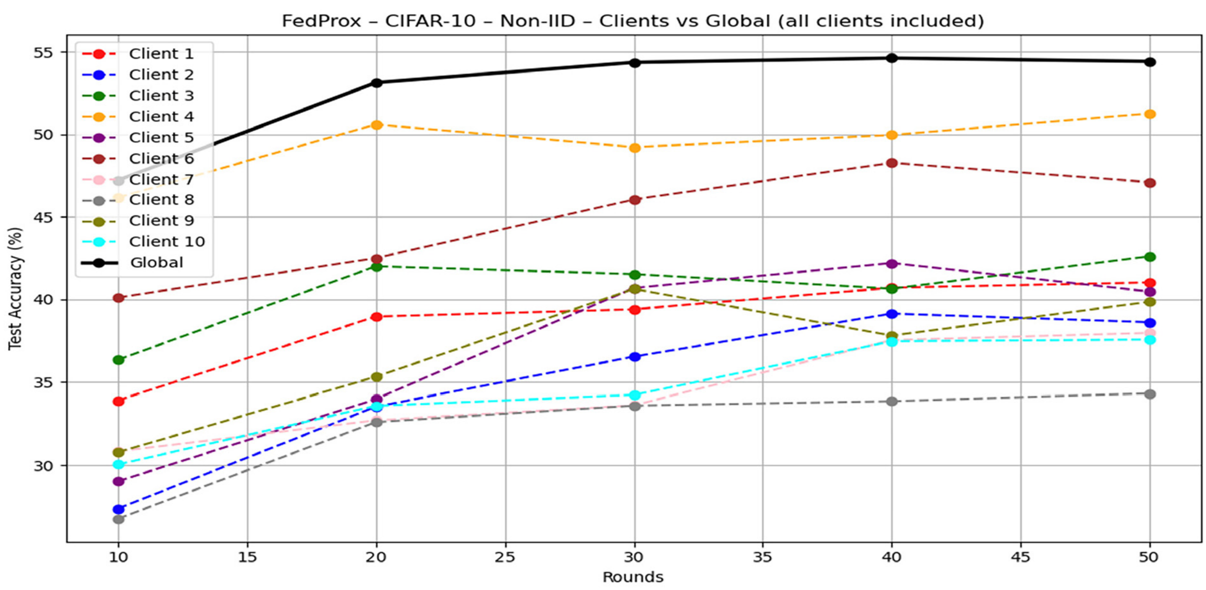

Figure 2. Client and global accuracy under Non-IID setting using FedProx (all clients included)

that certain clients have more representative or informative data, allowing them to dominate the learning process, while others struggle owing to limited or biased local distributions.

This imbalance highlights a key limitation of FL under heterogeneous conditions. Even when the global model performs reasonably well, its benefits are not evenly distributed across all clients. Clients with restricted or skewed data distributions, referred to as stragglers, tend to underperform, whereas dominant clients contribute disproportionately to the global update. As a result, the system exhibits uneven learning outcomes, raising concerns about fairness and equitable participation.

To further investigate this behavior, we examined the impact of removing dominant clients from the training process, as explored in prior work. Figure 3 and 4 present the results obtained after excluding the top-performing client and the top two performing clients, respectively. As shown, removing dominant clients reduces the performance gap between the remaining participants, leading to more uniform client accuracies. However, this improvement in balance is accompanied by a reduction in global model accuracy, indicating a trade-off between performance and fairness.

These findings demonstrate that client imbalance in FL is primarily driven by data heterogeneity. While manual exclusion of dominant clients can mitigate imbalance, it is not a practical solution, as it requires external intervention and may discard valuable information. This limitation motivates the need for an automated mechanism that can regulate client contributions dynamically without removing participants from the training process.

3.3. Adaptive Proximal Regulation (Initial Approach).

To address the limitations of manual client exclusion and enable automated fairness control, we introduce an adaptive proximal regulation mechanism within the FedProx framework. Instead of assigning a fixed proximal coefficient to all clients, the proposed approach dynamically adjusts the proximal regularization term based on client performance at each communication round.

The key concept is to regulate client contributions by modifying the strength of the proximal constraint according to how each client performs relative to others. At the beginning of each round, the server evaluates the accuracy of participating clients using their performance from the previous round. Let a_k denote the accuracy of client k, and $\grave{a}$ represent the mean accuracy across all selected clients.

In the initial approach, we hypothesize that **high-performing (dominant) clients** contribute disproportionately to the global model and may bias it toward their local data distributions. To mitigate this effect, we increase the proximal regularization term for such clients, forcing their updates to remain closer to the global model. The adaptive proximal coefficient is computed using Eq. 6.

$$\mu_k = \mu \left(1 + \frac{a_k - \grave{a}}{\grave{a} + \epsilon}\right) \quad \text{for } a_k > \grave{a} \tag{6}$$

where μ is the base proximal coefficient and ϵ is a small constant for numerical stability. For clients whose performance does not exceed the average, the base proximal coefficient is retained.

This formulation ensures that clients with higher-than-average accuracy are subject to stronger regularization, reducing their influence on the global update. The goal is to suppress excessive dominance and promote a more balanced contribution across clients without explicitly removing any participants.

To evaluate the effectiveness of this approach, we analyze both global model performance and fairness across clients. Fairness is measured using the standard deviation of client accuracies, where lower variance indicates more uniform performance. Figure 5 presents the evolution of global and client accuracies, along with the corresponding fairness metric under the adaptive proximal regulation strategy.

As observed, while the global model continues to converge, the adaptive mechanism does not consistently reduce performance disparity across clients. In some cases, the variation in client accuracies remains similar to or higher than that of the baseline, suggesting that penalizing dominant clients alone is insufficient for achieving fairness under heterogeneous data conditions.

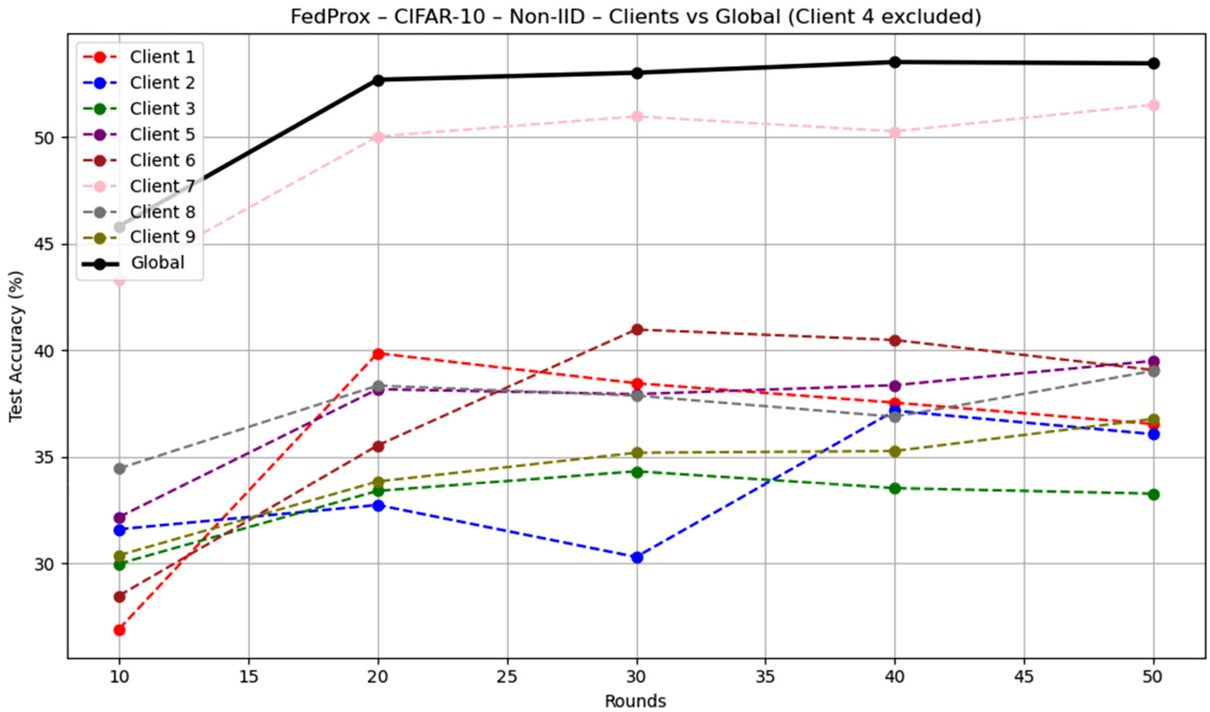

Figure 3. Client and global accuracy after removing the top-performing client

These results indicate that the relationship between client performance and fairness is more complex than initially assumed. Low-performing clients may exhibit strong local bias owing to limited or skewed data, which is not addressed by simply restricting dominant clients. This observation motivates a deeper investigation into the role of client behavior in FL and leads to the development of an improved adaptive strategy, discussed in the following section.

Table 1 shows that FedAvg and Adaptive FedProx achieve comparable global accuracy under IID conditions. However, the adaptive approach produces a higher fairness standard deviation, indicating that adaptive proximal regulation does not provide a fairness benefit when client data distributions are already balanced.

3.4. Failure Analysis under Non-IID. Though the adaptive proximal regulation strategy shows stable convergence under

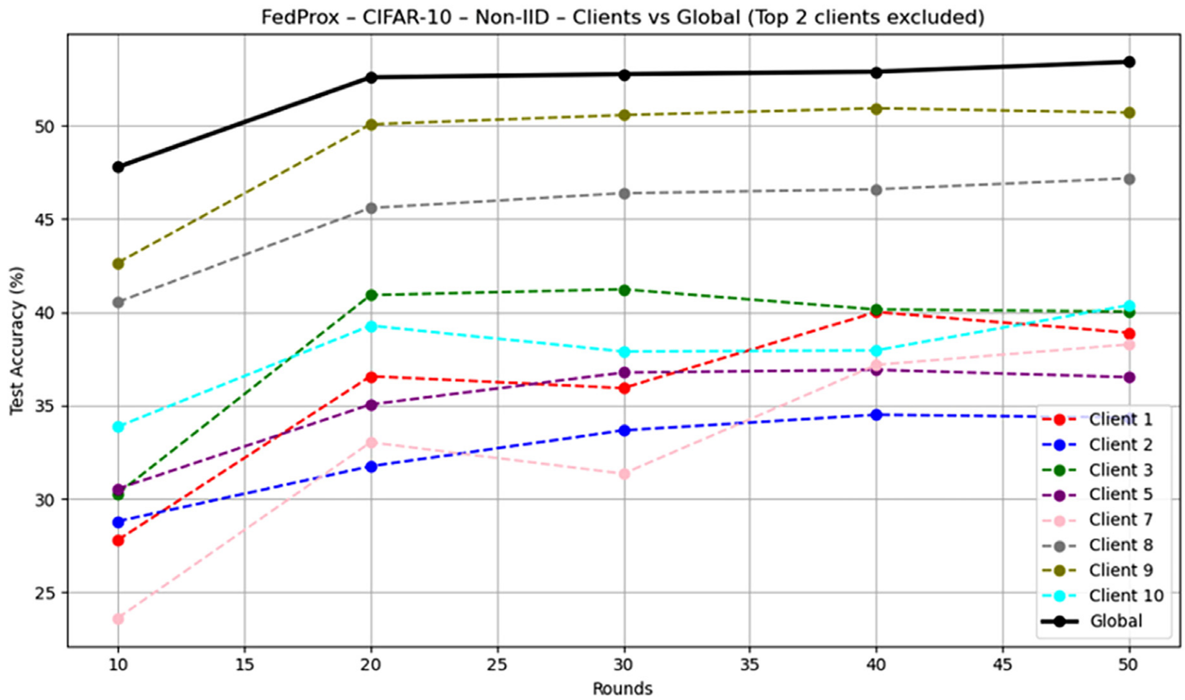

Figure 4. Client and global accuracy after removing the top two performing clients

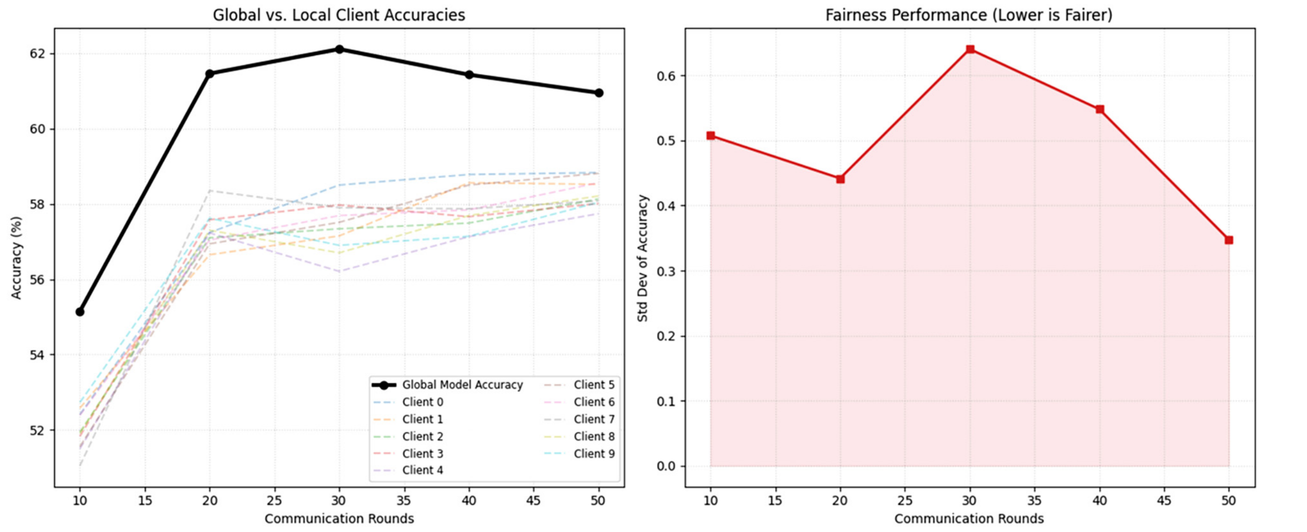

Figure 5. Global and client accuracy, along with fairness (standard deviation), under adaptive proximal regulation (initial approach)

IID conditions, its effectiveness significantly degrades in heterogeneous environments. To evaluate this behavior, we analyze the performance of the adaptive method under Non-IID data distributions, where client data is highly skewed and varies substantially across participants.

Figure 6 presents a comparison between the baseline FedAvg algorithm and adaptive proximal regulation approach under Non-IID conditions. As observed, although both methods achieve comparable levels of global accuracy, the adaptive approach leads to a substantially higher variation in client accuracies. This is reflected in the fairness metric, where the standard deviation of client performance is noticeably larger compared to that of the baseline.

This result indicates that the initial adaptive strategy[5,7] , which penalizes high-performing clients, is ineffective in reducing client imbalance under Non-IID settings. The approach can exacerbate the disparity between clients, leading to poorer fairness outcomes. This behavior reveals an important limitation in the underlying assumption of the method.

A closer examination of client behavior provides further insight into this phenomenon. In highly heterogeneous environments, low-performing clients often possess narrowly distributed or biased datasets. These clients tend to overfit their local data and deviate significantly from the global objective. By contrast, high-performing clients typically have access to more representative data distributions, enabling them to generalize better and contribute positively to the global model.

By increasing the proximal regularization term for high-performing clients, the adaptive strategy restricts their ability to influence the global update. At the same time, low-performing clients are allowed to train with relatively weaker constraints, which enables them to drift further toward their biased local optima. As a result, the gap between client performances widens, rather than narrowing as intended.

This observation highlights a critical insight: **client performance in FL does not directly correspond to fairness contribution**. High-performing clients are not necessarily the source of imbalance; instead, low-performing clients may introduce instability due to their biased local updates. Therefore, simply suppressing dominant clients is insufficient and may lead to unintended consequences.

These findings demonstrate that fairness-aware regulation in FL must account for the nature of client data distributions, rather than relying solely on performance-based suppression. This motivated the development of an alternative strategy that focuses on stabilizing low-performing clients, which is explored in the following section.

Table 2 shows that the initial adaptive proximal regulation strategy degrades both fairness and global accuracy under Non-IID conditions. Compared to the FedAvg baseline, Adaptive FedProx exhibits a significantly higher standard deviation in client accuracies, confirming that penalizing dominant clients increases performance disparity rather than reducing it in heterogeneous environments.

3.5. Reverse Adaptive Proximal Regulation.

Based on the limitations observed in the initial adaptive proximal regulation strategy, we reformulate an approach that reconsiders the role of client performance in heterogeneous environments. The previous results demonstrate that penalizing high-performing clients is ineffective, as it restricts clients that contribute positively to global learning while allowing low-performing clients to diverge further due to weak regularization.

Table 1. *IID performance comparison between FedAvg and Adaptive FedProx*

Method	Final Global Accuracy	Final Fairness (Std Dev)
FedAvg	60.86%	0.1850
Adaptive FedProx	60.95%	0.3474

Table 2. *Non-IID performance comparison between FedAvg and Adaptive FedProx*

Method	Final Global Accuracy	Final Fairness (Std Dev)
FedAvg	60.86%	0.4098
Adaptive FedProx	58.83%	0.8677

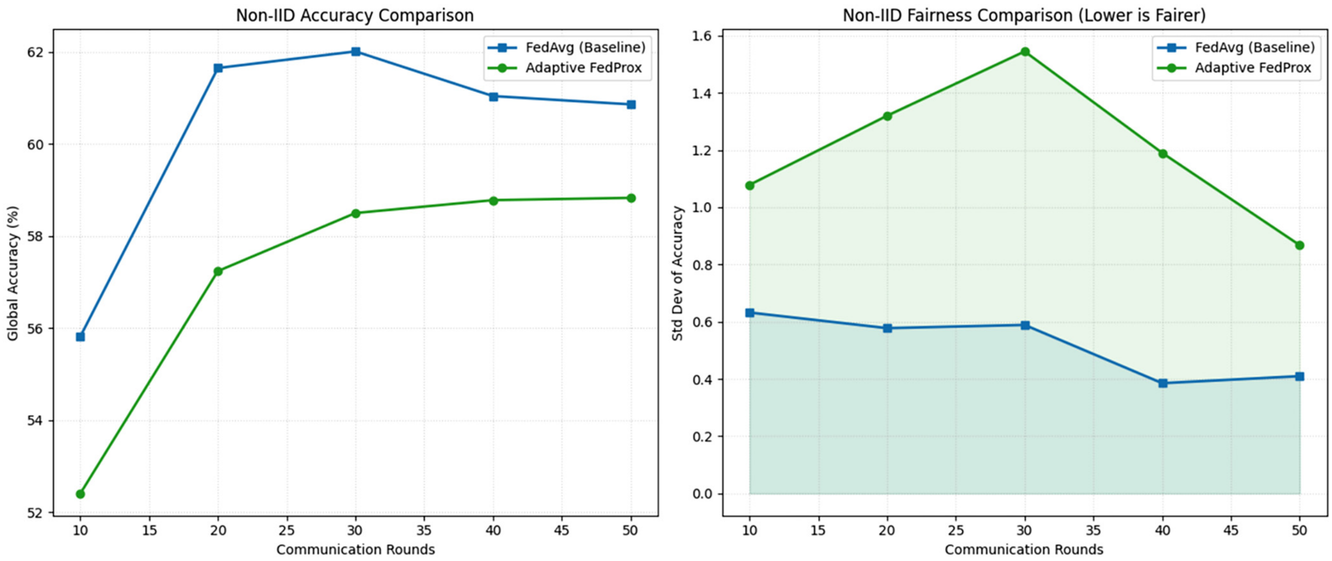

Figure 6. Comparison of global accuracy and fairness between FedAvg and reverse adaptive proximal regulation under Non-IID conditions

To address this, we propose a **reverse adaptive proximal regulation strategy**, in which stronger regularization is applied to low-performing clients (*stragglers*), whereas high-performing clients (*leaders*) are allowed greater flexibility during training. This design is motivated by the observation that straggler clients often exhibit strong local bias due to limited or skewed data distributions, causing them to deviate significantly from the global model.

In this approach, the proximal coefficient is dynamically adjusted such that clients with accuracy below the mean are assigned a higher proximal penalty, effectively anchoring their updates closer to the global model. Conversely, clients with above-average performance retain the base proximal coefficient, enabling them to contribute more freely to global aggregation. This selective regulation reduces excessive local drift from biased clients while preserving the beneficial influence of well-performing clients.

Figure 7 presents the performance of the reverse adaptive proximal regulation strategy under Non-IID conditions, compared with the baseline FedAvg algorithm. As observed, the proposed method achieves comparable or slightly improved global accuracy relative to the baseline. More importantly, the fairness metric shows a noticeable improvement during the early stages of training, indicating that the adaptive anchoring of low-performing clients helps reduce performance disparity across participants. However, it is also observed that the fairness improvement is not consistently maintained throughout all communication rounds. In later stages of training, the performance gap between clients slightly increases again. This behavior can be attributed to the constant strength of the proximal penalty, which may overly constrain straggler clients and limit their ability to adapt as the global model converges.

Despite this limitation, the reverse adaptive strategy provides a significant improvement over the initial approach by correctly identifying the source of imbalance and applying targeted regulation. These findings highlight that **fairness-aware proximal control in FL should prioritize stabilizing low-performing clients rather than suppressing dominant ones**.

This insight opens the door for further refinement of the method, such as incorporating a decaying proximal factor that gradually reduces the strength of regularization over time. Such an extension could preserve early-stage fairness benefits while avoiding late-stage stagnation, representing a promising direction for future work.

Table 3 shows that the proposed Reverse Adaptive FedProx strategy achieves slightly higher global accuracy compared to the FedAvg baseline while maintaining competitive fairness performance. These results indicate that applying stronger regularization to low-performing clients is a more effective fairness-aware strategy than suppressing dominant clients under Non-IID conditions.

4. RESULTS AND DISCUSSION

The experimental results in Section 3 provide a comprehensive evaluation of FL performance under IID and Non-IID settings, as well as the impact of adaptive proximal regulation strategies on fairness and global accuracy. In this section, we synthesize these findings and discuss their implications in the context of fairness-aware federated optimization.

Under IID conditions, the results demonstrate that FL naturally achieves balanced performance across clients. As shown in Section 3.1, all clients exhibit similar accuracy trends, and the variance in performance remains minimal. This confirms that fairness is not a significant concern when data is uniformly distributed, and additional regulation mechanisms are unnecessary in such settings. The adaptive proximal approach also shows

Table 3. *Non-IID performance comparison between FedAvg and Reverse Adaptive FedProx*

Method	Final Global Accuracy	Final Fairness (Std Dev)
FedAvg	60.86%	0.4098
Reverse Adaptive FedProx	61.04%	0.4611

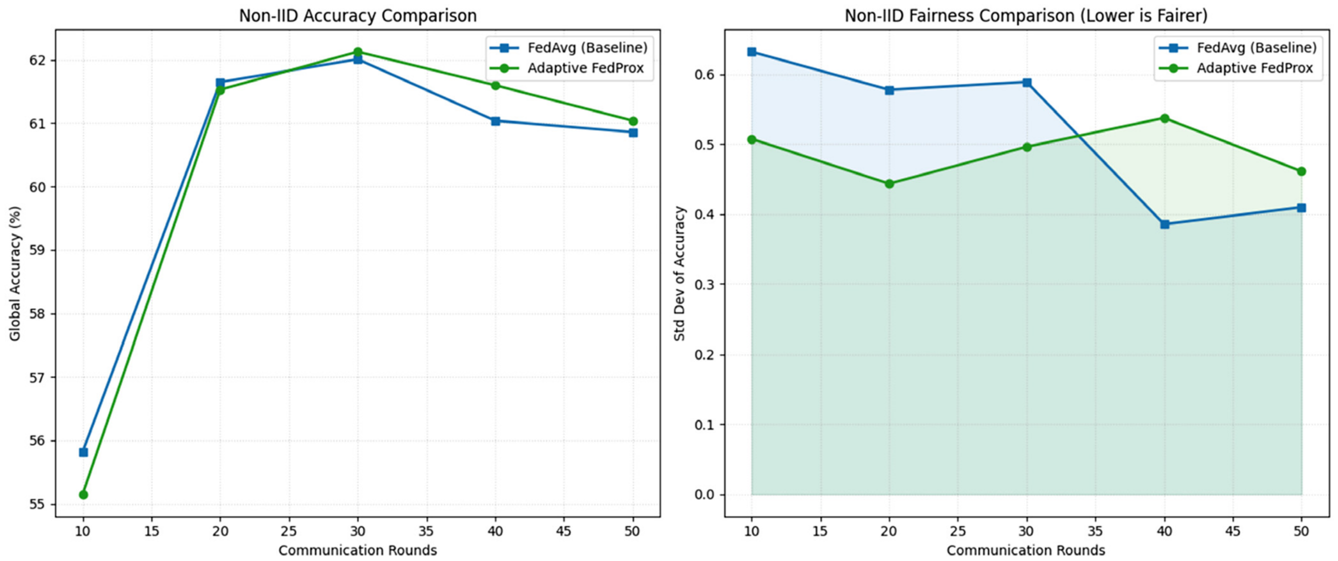

Figure 7. Comparison of global accuracy and fairness between FedAvg and adaptive proximal regulation under Non-IID conditions

no meaningful advantage under IID conditions, as the system is already stable and balanced.

By contrast, the Non-IID experiments reveal substantial performance disparities[6] across clients, highlighting the inherent challenge of client imbalance in FL. As discussed in Section 3.2, some clients consistently outperform others owing to more representative data distributions, while straggler clients struggle owing to limited or biased local data. The results further confirm that manual exclusion of dominant clients can reduce this imbalance but leads to a decline in global accuracy, illustrating a clear trade-off between fairness and performance.

The introduction of adaptive proximal regulation was intended to address this trade-off by dynamically controlling client contributions. However, the results from Sections 3.3 and 3.4 show that the initial strategy of penalizing dominant clients does not achieve the desired outcome. Instead, it increases performance disparity under Non-IID conditions, as reflected by a higher standard deviation of client accuracies. This finding highlights a critical misconception: dominant clients are not the primary cause of unfairness, and restricting their influence can negatively impact the overall learning process.

A deeper analysis revealed that low-performing clients play a more significant role in driving imbalance. These clients often overfit to their narrow local datasets, causing their updates to diverge from the global objective. When such clients are allowed to train with weak constraints, their biased updates amplify performance gaps rather than reducing them. This explains the reason behind failure of the initial adaptive strategy in improving fairness.

The reverse adaptive proximal regulation strategy proposed in Section 3.5 addresses this issue by shifting the focus from suppressing dominant clients to stabilizing straggler clients. By applying stronger regularization to low-performing clients, the method reduces excessive local drift and aligns their updates more closely with the global model. The results show that this approach improves fairness during the early stages of training while maintaining competitive global accuracy.

Nevertheless, the analysis also reveals the limitations of the proposed method. The fairness improvement is not consistently sustained across all training rounds, as the fixed strength of the adaptive penalty may overly constrain straggler clients in later stages. This suggests that a more refined approach, such as incorporating a time-dependent or decaying proximal factor, may further enhance both fairness and convergence.

Overall, the results demonstrate that fairness-aware FL requires a nuanced understanding of client behavior under heterogeneous data conditions. Rather than treating all clients symmetrically or focusing solely on high-performing participants, effective regulation must account for the dynamics of low-performing clients and their impact on global learning. The proposed reverse adaptive strategy represents a meaningful step toward automated fairness control, while also highlighting important directions for future research.

5. CONCLUSIONS

This study investigated fairness challenges[2,7] in FL, with a particular focus on the impact of client heterogeneity under Non-IID data distributions. Although FL enables decentralized model training with strong privacy guarantees, it often leads to imbalanced performance across clients owing to differences in local data distributions. Addressing this imbalance is essential to ensure equitable participation and reliable model performance in real-world applications.

Initial experiments confirmed that, under IID conditions, FL achieves naturally balanced performance across clients, and fairness is not a significant concern. However, under Non-IID settings, substantial disparities emerge between clients, with some achieving significantly higher accuracy than others. This client imbalance highlights a key limitation of standard FL approaches and motivates the need for fairness-aware mechanisms.

To address this issue, this study proposed an adaptive proximal regulation framework based on the FedProx algorithm. The initial approach aimed to improve fairness by increasing the proximal regularization term for high-performing clients. However, experimental results demonstrated that this strategy does not effectively reduce performance disparity and can, in some cases, worsen fairness under heterogeneous conditions. This finding reveals that suppressing dominant clients is insufficient for achieving equitable outcomes in FL.

A deeper analysis showed that low-performing clients, often characterized by highly biased local datasets, contribute significantly to performance imbalance. Based on this insight, a reverse adaptive proximal strategy was introduced, in which stronger regularization was applied to low-performing clients while allowing high-performing clients greater flexibility. This approach improves fairness during the early stages of training and maintains competitive global accuracy, demonstrating a more effective direction for fairness-aware federated optimization.

Despite these improvements, the results also indicate limitations in consistently maintaining fairness across all training rounds. Using a fixed adaptive scaling factor may overly constrain certain clients in later stages, suggesting the need for further refinement. Future work will explore extensions such as time-dependent or decaying proximal regulation, as well as more sophisticated fairness metrics and adaptive strategies that better capture client behavior.

Overall, this study highlights that fairness in FL is not solely determined by dominant clients, but is strongly influenced by the behavior of low-performing clients under heterogeneous data conditions. The proposed adaptive proximal regulation framework provides a step toward automated fairness control and offers valuable insights for designing more robust and equitable FL systems.

AFFILIATIONS AND AUTHOR DETAILS

Undergraduate Author

Faris Misfir Alshahrany – *Department of Information and Computer Science, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*; 0009-0006-1492-7327
Email: s202277660@kfupm.edu.sa;

Corresponding Author

Hamoud Ibrahim Aljamaan – *Research Mentor, Professor, Department of Information and Computer Science, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*; 0000-0002-2146-9348
Email: hjamaan@kfupm.edu.sa

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the Department of Information and Computer Science at King Fahd University of Petroleum and Minerals (KFUPM) for providing the academic environment and resources necessary for this research.

Special thanks are extended to Dr. Hamoud Ibrahim Aljamaan for their guidance, support, and valuable insights throughout the development of this work. Their expertise and mentorship have been instrumental in shaping the direction and quality of this research.

The authors also appreciate the support provided by the university facilities and computing resources that enabled the successful completion of the experimental work.

REFERENCES

(1) McMahan, H. B. et al. Communication-efficient learning of deep networks from decentralized data. *Proc. AISTATS* **54**, 1273–1282 (2017).

(2) Li, T., Sahu, A. K., Talwalkar, A. & Smith, V. Federated optimization in heterogeneous networks. *Proc. MLSys* **2**, 429–450 (2020).

(3) Kairouz, P. et al. Advances and open problems in federated learning. *Found. Trends Mach. Learn.* **14**, 1–210 (2021).

(4) Mohri, M., Sivek, G. & Suresh, A. T. Agnostic federated learning. *Proc. ICML*, 4615–4625 (2019).

(5) Karimireddy, S. P. et al. SCAFFOLD: Stochastic controlled averaging for federated learning. *Proc. ICML*, 5132–5143 (2020).

(6) Zhao, Y. et al. Federated learning with non-IID data. *arXiv preprint* arXiv:1806.00582 (2018).

(7) Nishio, T. & Yonetani, R. Client selection for federated learning with heterogeneous resources. *IEEE ICC*, 1–7 (2019).

(8) Cho, Y. J., Park, J., Bennis, M. & Kim, S.-L. Client selection in federated learning for resource-constrained devices. *IEEE Trans. Wireless Commun.* **19**, 7108–7122 (2020).

(9) Duan, M. et al. Adaptive client sampling for fair federated learning. *Proc. AAAI*, 9462–9470 (2021).

(10) Huang, L., Yang, Y. & Han, S. Fairness-aware federated learning via dynamic client participation. *NeurIPS Workshop* (2020).

(11) Zeng, Y. et al. Fairness in federated learning: Challenges and opportunities. *Proc. IJCAI*, 4703–4710 (2021).

(12) Lim, W. et al. Federated learning in mobile edge networks: A survey. *IEEE Commun. Surv. Tutorials* **22**, 2031–2063 (2020).