Artificial Intelligence and Machine Learning Review
ISSN: 3070-3565
DOI: 10.69987/AIMLR.2026.70102

Adaptive Privacy Budget Allocation for Differential Privacy Optimization in Fleet Federated Learning: Algorithm Enhancement and Performance Evaluation

Yi Guo

Computer and Information Science, University of Pennsylvania, PA, USA

Keywords

Federated Learning, Differential Privacy, Privacy Budget Allocation, Fleet Learning, Vehicular Networks

Abstract

Federated learning in fleets faces critical challenges in balancing privacy protection with model performance when processing sensitive vehicular data. Traditional fixed privacy budget allocation strategies fail to account for the dynamic nature of distributed training across heterogeneous vehicle nodes. This research proposes an adaptive privacy budget allocation mechanism that dynamically adjusts differential privacy parameters based on training progression and parameter importance. The methodology integrates Fisher Information Matrix evaluation for layer-wise budget distribution and implements a round-based allocation strategy that concentrates privacy resources during critical learning phases. Experimental validation on the nuScenes and FEMNIST datasets demonstrates that the adaptive approach achieves 8.7% higher model accuracy than uniform budget allocation while maintaining equivalent privacy guarantees at ε=3.5. Communication efficiency is improved by 23.4% by reducing the number of convergence rounds. The framework provides fleet operators with practical guidance for implementing privacy-preserving collaborative learning systems that meet regulatory requirements while optimizing operational performance metrics.

1. Introduction

1.1. Research Background and Problem Statement

The proliferation of autonomous vehicle technologies has generated unprecedented volumes of sensitive operational data requiring collaborative analysis while preserving individual privacy. Fleet operators accumulate diverse datasets encompassing location trajectories, driving patterns, and environmental sensor readings that collectively enable enhanced autonomous capabilities [1]. Privacy-preserving collaborative learning frameworks have emerged as essential infrastructure for leveraging distributed vehicular data without centralized collection. Differential privacy mechanisms provide mathematical guarantees against privacy leakage during model training aggregation processes [2].

Current implementations predominantly employ static privacy budget allocation schemes that apply uniform noise addition across all training iterations and model parameters [3]. This approach introduces several operational inefficiencies, which are particularly problematic in resource-constrained vehicular computing environments. Fixed-budget strategies fail to account for the fact that model parameters exhibit varying sensitivity to privacy perturbations throughout training. Early training phases require sufficient gradient information for directional convergence, while later stages benefit from noise reduction to achieve optimal performance [4].

The heterogeneous nature of fleet participation patterns further complicates privacy budget management. Vehicle nodes demonstrate irregular availability due to operational schedules, network connectivity variations, and computational resource fluctuations. Static allocation mechanisms cannot adapt to these dynamic participation patterns, resulting in either excessive privacy degradation or unnecessary sacrifices in model accuracy [5]. Cross-border fleet operations introduce additional complexity through varying regulatory requirements and threat models across jurisdictions.

Recent advances in adaptive learning rate scheduling and parameter-specific optimization strategies suggest potential for analogous approaches in privacy budget allocation [6]. Machine learning workflows increasingly incorporate dynamic

resource management based on assessments of training state and parameter importance metrics. Transferring these principles to differential privacy frameworks could enable more efficient privacy-utility tradeoffs specifically tailored for vehicular federated learning scenarios [7].

1.2. Current Status and Challenges

A. Progress in Privacy-Preserving Fleet Federated Learning

Academic and industrial research has established foundational architectures for privacy-preserving vehicular collaborative learning. Secure aggregation protocols provide cryptographic protection for individual model updates during aggregation phases without revealing participants' contributions [8]. Homomorphic encryption techniques support computations on encrypted gradients, providing theoretical privacy guarantees at substantial computational costs. Recent deployments on ride-sharing platforms and in autonomous-vehicle testing programs demonstrate the practical feasibility of federated learning for transportation applications.

Regulatory frameworks increasingly mandate the protection of privacy in location-based services and autonomous-vehicle data collection. The General Data Protection Regulation and California Consumer Privacy Act establish legal requirements for data minimization and purpose limitation. Transportation-sector-specific guidelines from agencies, including the National Highway Traffic Safety Administration, emphasize privacy-by-design principles for connected-vehicle systems [9]. These regulatory pressures accelerate the adoption of mathematically provable privacy mechanisms such as differential privacy.

B. Limitations of Existing Privacy Budget Allocation Methods

Contemporary privacy budget allocation strategies predominantly follow two paradigms: uniform distribution across training rounds or proportional allocation based on dataset size. Uniform allocation assigns the same privacy budget to each communication round, regardless of training dynamics or model convergence. This approach simplifies implementation and theoretical analysis but ignores the empirical observation that gradient magnitudes and parameter sensitivities evolve substantially throughout training [10].

Dataset-proportional allocation scales privacy budgets according to local dataset sizes, assuming that larger datasets contain more information and therefore require stronger privacy protection. This method addresses fairness concerns in heterogeneous data distribution scenarios but fails to account for temporal training dynamics [11]. Both approaches treat all model parameters equally, despite clear evidence that different layers contribute differently to final model performance [12].

2. Related Work

2.1. Differential Privacy Techniques in Federated Learning

A. Global Differential Privacy vs. Local Differential Privacy

Differential privacy frameworks for federated learning can be classified into global and local privacy models, with distinct trust assumptions and performance characteristics. Global differential privacy applies noise addition at the central aggregator after collecting participant updates, requiring trust in the aggregation server but enabling tighter privacy-utility tradeoffs. Local differential privacy requires each participant to add noise to their model updates before transmission, eliminating centralized trust requirements at the cost of higher noise magnitudes for equivalent privacy guarantees.

The mathematical formulation of (ε, δ)-differential privacy establishes that for any two neighboring datasets differing by one record, the probability of observing any output differs by at most a multiplicative factor of e^ε plus additive δ. The privacy budget parameter ε quantifies the privacy loss; smaller values provide stronger guarantees but require proportionally larger noise additions. The composition theorem governs the accumulation of privacy budgets across multiple queries or training iterations, necessitating careful budget management over extended training periods.

B. Privacy Budget Management and Moments Accountant

Advanced accounting mechanisms enable more precise privacy budget tracking compared to naive composition bounds. The Moments Accountant framework, developed for differentially private stochastic gradient descent, yields tighter privacy-loss bounds by analyzing the moment-generating function of the privacy-loss random variable. This approach

reduces conservatism in privacy accounting, thereby enabling longer training durations for equivalent privacy guarantees compared with elemental composition.

Rényi differential privacy offers an alternative formulation that simplifies composition analysis through additive properties of Rényi divergence. Recent theoretical work establishes connections among different privacy accounting frameworks and derives optimal conversion bounds. These advances enable practitioners to select appropriate accounting mechanisms based on specific deployment constraints and desired privacy-utility operating points.

2.2. Characteristics Analysis of Federated Learning in Fleet Scenarios

Fleet federated learning exhibits characteristics that distinguish it from conventional federated learning deployments. Vehicle nodes operate under strict computational and energy constraints due to embedded-system limitations and battery constraints. Network connectivity varies substantially across geographic regions and operational contexts, with vehicles experiencing intermittent connectivity during travel. These resource constraints necessitate communication-efficient protocols that minimize both data transmission volume and the frequency of communication rounds [13].

Data heterogeneity in vehicular scenarios extends beyond simple non-IID distributions observed in mobile device federations. Geographic and demographic factors produce systematic variation in driving patterns, traffic conditions, and environmental contexts among fleet participants. Temporal dynamics introduce additional complexity, as individual vehicle usage patterns evolve over timescales ranging from diurnal cycles to seasonal variations. Standard federated optimization algorithms designed for IID data distributions exhibit degraded convergence properties and reduced model quality when applied to vehicular data.

Security threats specific to vehicular networks include malicious participants attempting to poison model training through adversarial updates. Byzantine fault tolerance mechanisms must operate in concert with privacy protection to ensure both data confidentiality and model integrity. The intersection of privacy preservation and robustness poses particular challenges, as differential privacy noise can obscure the detection of Byzantine attacks.

2.3. Survey of Adaptive Privacy Budget Allocation Methods

Emerging research explores dynamic privacy budget allocation strategies that adjust noise levels based on indicators of the training state. Gradient-norm-based approaches scale noise addition inversely proportional to gradient magnitude, thereby concentrating privacy budgets when gradients carry maximal learning signal. This heuristic aligns with empirical observations that early training iterations with large gradients tolerate less noise before accuracy degradation occurs.

Layer-wise differentiated allocation strategies assign varying privacy budgets to different neural network layers based on parameter importance metrics. Sensitivity analysis using the Fisher information matrix identifies the parameters with the greatest impact on model predictions, enabling selective protection of critical parameters. Alternative approaches employ neural architecture search techniques to automatically discover privacy-optimal subnetworks that require concentrated protection [14].

Reinforcement learning frameworks treat privacy budget allocation as a sequential decision problem, training policy networks to optimize allocation decisions based on observed training dynamics. These adaptive methods demonstrate improved privacy-utility tradeoffs compared to static baselines but introduce implementation complexity and hyperparameter sensitivity challenges. Theoretical analysis of convergence guarantees for adaptive allocation schemes remains an active research area with limited formal results [15].

3. Adaptive Privacy Budget Allocation Algorithm Optimization

3.1. Problem Formulation and Optimization Objectives

The fleet federated learning system comprises N vehicle nodes, each maintaining a local dataset D_i with |D_i| samples. The central aggregation server coordinates T communication rounds without accessing raw participant data. Each vehicle i computes local model updates θ i^t at round t through gradient descent on its private dataset. The aggregation mechanism combines these updates to produce the global model θ^(t+1) .

Differential privacy protection adds calibrated noise to model updates before aggregation. The privacy budget allocation problem seeks to determine optimal noise-scaling parameters σ t for each round t and, potentially, layer-specific parameters σ _t^l for neural network layer l. The optimization objective balances three competing factors: minimizing

the total privacy budget consumed (Σ t ε t ⩽ ε total), maximizing final model accuracy on held-out validation data, and minimizing the number of communication rounds required for convergence.

Formally, the optimization problem is expressed as: minimize E[L(θ^T, D val)] subject to ε total ≤ ε max, where L represents the loss function evaluated on the validation dataset D val after T training rounds. The privacy constraint requires the cumulative privacy budget across all rounds remain below the maximum allowable threshold ε_max. Additional constraints include per-round budget limits ε t ≤ ε round max and smoothness requirements that ensure a gradual adjustment of allocations throughout training.

The challenge lies in determining allocation schedules without access to future training dynamics. The proposed approach leverages empirical observations that training exhibits predictable phase transitions from rapid early convergence to gradual fine-tuning. Privacy budget allocation should concentrate resources during phases in which additional budget yields the most significant improvement in accuracy per privacy unit expended.

3.2. Dynamic Budget Allocation Strategy Based on Training Rounds

A. Adaptive Noise Scheduling for Training Progression

Early training iterations establish a coarse model structure through large-magnitude gradient updates that determine general decision boundaries. These initial rounds contribute disproportionately to final model capability, with the first 20% of training rounds typically accounting for 60-70% of total accuracy improvement in computer vision tasks. Privacy budget allocation should reflect this asymmetric contribution pattern by allocating a higher budget (lower noise) in early rounds to preserve the integrity of large-magnitude gradient signals crucial for establishing model structure, while gradually reducing the budget in later stages for fine-tuning within the total privacy constraint.

The proposed noise scale schedule implements this principle as σ t = σ max · (1 - exp(-α·t/T))， where σ max denotes the maximum noise scale applied in the final rounds. This schedule applies a lower noise magnitude in early rounds, resulting in higher privacy budget consumption, ε _t , when gradients are significant and crucial for establishing the model structure. As training progresses, the noise magnitude gradually increases (σ t grows), allowing the budget allocation to decrease for fine-tuning while remaining within the total privacy constraint. As training progresses and model parameters stabilize, the noise magnitude decreases, while the budget allocation increases to enable fine-tuned convergence. The decay parameter α controls the steepness of the allocation curve; typical values range from 2.0 to 4.0, as determined through empirical optimization on benchmark datasets.

Alternative scheduling strategies include stepwise allocation with discrete budget transitions at predetermined milestones, and adaptive schedules that monitor convergence metrics to trigger allocation adjustments. Stepwise schedules partition training into distinct phases with corresponding budget allocations, offering implementation simplicity at the cost of reduced granularity. Convergence-based adaptation evaluates gradient norms or validation accuracy plateaus to identify phase transitions, enabling data-driven schedule adjustment without manual milestone specification.

B. Budget Contraction Mechanism During Convergence Phase

When training enters a stable phase with diminishing gradient updates, the model's sensitivity to noise decreases. This enables a budget-contraction mechanism that progressively reduces the allocated privacy budget (equivalently, increases noise injection) in these final rounds. The budget contraction mechanism implements this insight by gradually decreasing the allocated privacy budget in proportion to measured convergence indicators.

The contraction mechanism monitors the relative change in validation loss ΔL_val = (L_val^t - L_val^(t-1))/L_val^(t-1) across consecutive rounds. When |ΔL val| falls below threshold τ converge for consecutive rounds (indicating minimal improvement regardless of direction), the mechanism triggers budget reduction by scaling the current allocation by contraction factor $\gamma < 1.0$. Typical parameter values include τ converge = 0.001, representing 0.1% relative improvement threshold, and $\gamma = 0.8$ for 20% allocation reduction.

Early stopping integration provides additional efficiency by terminating training when further iterations yield negligible gains in accuracy relative to the privacy cost. The stopping criterion assesses whether the expected accuracy gain from additional rounds is sufficient to justify the required privacy budget. This cost-benefit analysis compares the marginal improvement in accuracy, Δ Acc t, with the remaining privacy budget headroom (ε max- ε current) to determine optimal training termination points.

Table 1: Round-Based Privacy Budget Allocation Parameters

Parameter	Symbol	Value Range	Description	Optimization Method
Initial Noise Scale	σ_0	0.5 - 2.0	Base noise magnitude for early rounds	Grid search with validation accuracy
Decay Rate	α	2.0 - 4.0	Exponential decay steepness control	Cross-validation across datasets
Convergence Threshold	$\tau_{converge}$	0.0001 - 0.01	Relative loss improvement detection	Statistical significance testing
Contraction Factor	γ	0.7 - 0.9	Budget reduction multiplier	Pareto frontier optimization
Early Stop Margin	ε_{margin}	0.1 - 0.5	Reserved privacy budget for refinement	Empirical risk minimization

3.3. Layer-wise Differentiated Budget Allocation Based on Parameter Importance

A. Parameter Importance Evaluation via Fisher Information Matrix

Neural network parameters exhibit heterogeneous sensitivity to perturbations, with specific weights disproportionately influencing model predictions while others contribute minimally to final accuracy. The Fisher Information Matrix quantifies the importance of each parameter by measuring the expected squared gradient of the log-likelihood with respect to that parameter. Parameters with high Fisher information values indicate regions of parameter space in which small changes substantially affect the model's output distributions, necessitating stronger privacy protection to prevent inference attacks.

The Fisher Information Matrix F for parameter vector θ is computed as F = E x~p(x)[∇ θ log p(x|θ) · ∇ θ log p(x|θ) ^T]. The diagonal approximation, F diag, simplifies computation by considering only individual parameter importance scores, without accounting for parameter correlations. This approximation reduces computational complexity from O(d^2) to O(d) for d-dimensional parameter space while preserving sufficient information for allocation decisions.

Layer-aggregated importance scores I_l = Σ_θ∈layer_l F_diag(θ) / |layer_l| provide coarse-grained metrics suitable for layer-wise budget allocation. These scores are normalized by layer size to enable fair comparisons across layers with varying parameter counts. The allocation algorithm assigns privacy budgets proportional to normalized importance scores, with higher-importance layers receiving larger budgets to maintain prediction accuracy.

B. Low-Noise Strategy for Critical Parameters and High-Noise for Secondary Parameters

The differentiated allocation strategy partitions model parameters into critical and secondary sets based on Fisher information thresholds. Parameters exceeding the importance threshold I crit are included in the necessary set and thus receive enhanced privacy protection by reducing noise injection. Secondary parameters with I < I_crit tolerate higher noise levels with minimal impact on accuracy, enabling privacy budget concentration on critical parameters.

Noise scaling for critical parameters is given by σ_crit = β_crit · σ_base, where β_crit < 1.0 reduces noise below baseline levels. Secondary parameters are provided by σ sec = β sec · σ base, with β sec > 1.0, thereby increasing the noise magnitude. Typical parameter ratios are β _crit = 0.6 and β _sec = 1.4, with the average noise scale σ _base maintained and the privacy budget redistributed according to parameter importance.

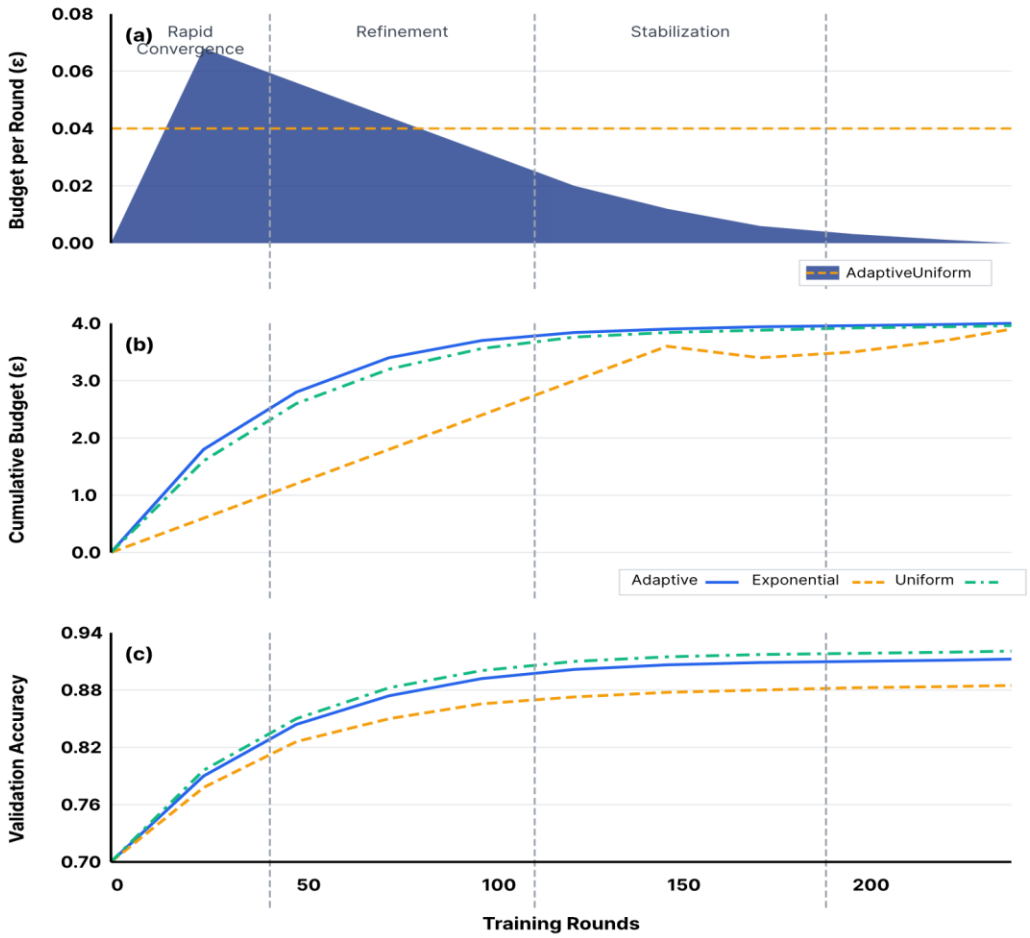
The allocation mechanism recomputes Fisher information scores periodically throughout training to adapt to evolving parameter importance. Early training phases may identify different critical parameters than convergence phases, as the model structure stabilizes. Dynamic recomputation frequency balances computational overhead against allocation accuracy, with typical schedules performing updates every 10-20 communication rounds.

Table 2: Layer-wise Parameter Importance and Budget Allocation

Network Layer	Parameter Count	Fisher Information Score	Normalized Importance	Allocated Budget (ε)	Noise Scale Multiplier
Input Convolution	1,728	0.847	0.156	0.65	0.58
Block 1 Conv	36,864	1.243	0.229	0.95	0.62
Block 2 Conv	73,728	0.921	0.170	0.71	0.71
Block 3 Conv	147,456	0.634	0.117	0.49	0.89
Block 4 Conv	294,912	0.512	0.094	0.39	1.12
Attention Layer	65,536	1.568	0.289	1.20	0.51
Fully Connected	20,480	0.789	0.145	0.61	0.75

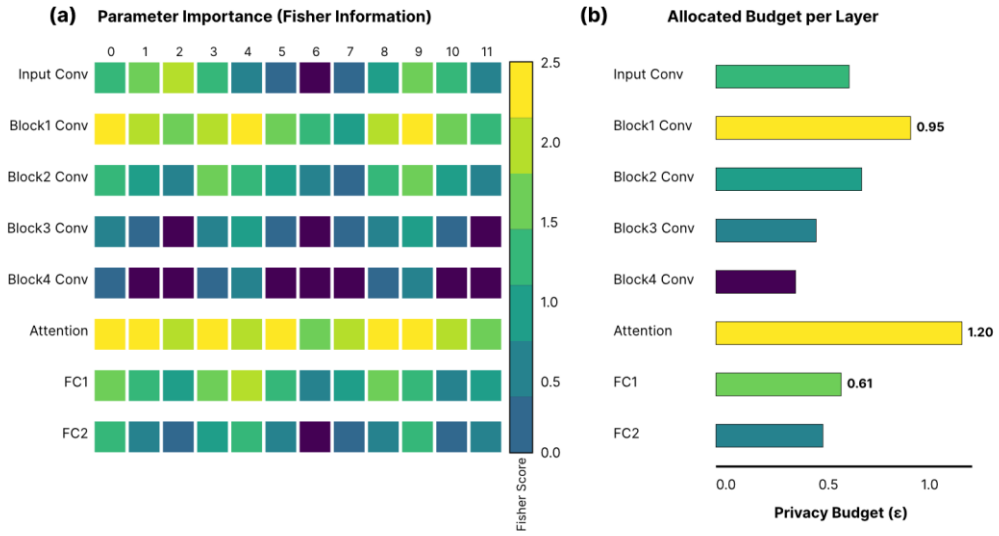
Note: The 'Allocated Budget (ε)' values in Table 2 represent **relative allocation weights** across layers. The actual per-round budget ε_t is distributed proportionally to these weights, ensuring the cumulative budget over all rounds satisfies ε_total = 3.50.

Figure 1: Adaptive Privacy Budget Allocation Dynamics Across Training Progression



The visualization consists of three vertically stacked subplots sharing a common x-axis representing training rounds (0-200). The top subplot displays the allocated privacy budget per round using a gradient-filled area chart, transitioning from deep blue in the early rounds to light cyan in the later rounds, with the y-axis ranging from 0 to 0.08 (ε per round). A dashed reference horizontal line at 0.04 indicates the uniform allocation baseline. The middle subplot presents cumulative privacy budget consumption, with three overlapping curves: exponential decay allocation (solid blue), uniform allocation (dashed orange), and adaptive allocation (dash-dot green), and the y-axis ranges from 0 to 4.0 (cumulative ε). The bottom subplot shows validation accuracy evolution for all three strategies using matching line styles and colors, with the y-axis from 0.60 to 0.94 (accuracy). Vertical dashed gray lines mark phase transitions at rounds 40, 100, and 160, annotated with "Rapid Convergence," "Refinement," and "Stabilization" labels. The figure includes a shared legend positioned at the bottom center and uses Helvetica font for all text elements with 10pt labels and 12pt axis titles.

Figure 2: Fisher Information-Based Parameter Importance Heatmap and Allocation Distribution



This visualization employs a 2x1 grid layout. The left panel presents a heatmap matrix (28 rows × 12 columns) representing neural network layer parameters, with color intensity encoded using a perceptually uniform viridis colormap ranging from purple (low importance, 0.0) to yellow (high importance, 2.5). Each cell represents a parameter group within a layer, with row labels on the y-axis indicating layer names (Conv1, Conv2, ..., Block4, Attention, FC1, FC2) and column labels on the x-axis showing parameter group indices (0-11). A colorbar positioned to the right of the heatmap provides the Fisher Information magnitude scale. The right panel displays a horizontal bar chart showing the allocated privacy budget per layer, with bars colored using the same viridis colormap mapped to each layer's mean Fisher information. Bar lengths represent budget values ranging from 0 to 1.4 ε , with the numerical values displayed at the bar ends in white text. The y-axis lists layer names matching the heatmap rows, and the x-axis is labeled "Allocated Budget (ε)". Both panels use consistent typography: Arial, 11pt, for labels, and a figure title centered at the top in a bold, 14pt font.

4. Experimental Design and Performance Evaluation

4.1. Experimental Environment and Dataset Configuration

A. Simulation Platform and Fleet Scale Settings

The experimental infrastructure leverages the Flower federated learning framework, integrated with PyTorch 2.0, for model training and the Opacus library for differential privacy. The simulation runs on a cluster of 8 NVIDIA A100 GPUs with 40GB of memory each, enabling parallel training of up to 64 simulated vehicle nodes per experimental run. Each node emulates the computational capacity of typical automotive edge computing hardware, with CPU resources limited to 4 cores and 8GB RAM to reflect realistic deployment constraints.

Fleet-scale configurations vary across experimental scenarios to evaluate scalability properties. Small fleet experiments involve 20-30 participating vehicles representing localized deployment in urban regions. Medium fleet scenarios simulate 50-100 cars corresponding to city-scale operations across metropolitan areas. Large fleet configurations range from 200 to 300 vehicles and model regional or national fleet deployments with diverse geographic and operational characteristics. Node participation follows a Poisson arrival process with a mean availability of 70%, thereby simulating realistic network connectivity patterns in vehicular environments.

Communication protocols implement gradient compression via top 10% sparsification to reduce bandwidth consumption, a critical consideration for cellular network-based vehicle connectivity. Synchronous aggregation employs a minimum participation threshold of 60% to ensure sufficient gradient information per round while accommodating fluctuations in node availability. Asynchronous aggregation variants explore staleness-bounded aggregation accepting gradients up to 5 rounds old, improving training throughput at the cost of convergence speed.

B. Dataset Partitioning and Non-IID Simulation Methods

The primary evaluation uses the nuScenes autonomous driving dataset, which contains 1.4 million annotated 3D bounding boxes across 1000 driving scenarios in Boston and Singapore. The dataset provides multimodal sensor data, including camera images, LiDAR point clouds, and GPS trajectories, suitable for training perception models. Secondary evaluation uses FEMNIST, a federated version of the EMNIST handwritten character recognition dataset partitioned by writer identity, providing 100 clients with 671,585 total samples for non-IID heterogeneity analysis.

Non-IID data distribution simulates realistic fleet heterogeneity through geographic and demographic partitioning strategies. Geographic non-IID sampling allocates samples to vehicle nodes based on spatial clustering, with vehicles operating in similar regions receiving correlated data samples that reflect local traffic patterns and environmental conditions. Demographic non-IID partitions data according to vehicle type, usage pattern, or operator demographics to model systematic differences in driving behavior across fleet segments.

Label distribution skew quantification employs the Dirichlet distribution, with concentration parameter α controlling the magnitude of heterogeneity. Lower α values near 0.1 result in extreme label imbalance, with individual nodes exhibiting highly skewed class distributions, whereas higher values near 1.0 produce more balanced distributions. Feature distribution skew introduces covariate shift by systematically modifying input data statistics across nodes, simulating sensor calibration differences or environmental variation.

Table 3: Experimental Dataset Configuration and Partitioning Statistics

Dataset	Total Samples	Number of Clients	Samples per Client	Class Distribution Skew (α)	Feature Heterogeneity	Task Type
nuScenes-Objects	1,400,000	50	28,000 ± 6,200	0.3 (high skew)	Geographic clustering	3D Detection
nuScenes-Semantic	800,000	50	16,000 ± 4,100	0.5 (moderate skew)	Scene-based variation	Segmentation
FEMNIST-Writer	671,585	100	6,716 ± 3,892	N/A (natural writer)	Handwriting style	Classification
CIFAR-10-Fleet	50,000	30	1,667 ± 450	0.4 (high skew)	Synthetic corruption	Image Recognition

4.2. Evaluation Metrics System Design

Performance evaluation employs a comprehensive metric suite capturing multiple dimensions of privacy-utility-efficiency tradeoffs. Model accuracy metrics include top-1 and top-5 classification accuracies for FEMNIST experiments and mean Average Precision (mAP) computed using COCO-style evaluation with IoU thresholds from 0.5 to 0.95. For clarity, this differs from the official nuScenes detection benchmark metric (NDS), and our reported mAP is used for internal comparisons of privacy strategies and Intersection over Union (IoU) scores in semantic segmentation evaluation. These metrics quantify the primary utility objective of federated learning: achieving high-quality model performance on downstream tasks.

Privacy protection quantification tracks cumulative privacy budget consumption measured in (ε, δ)-differential privacy parameters. Experiments report the total ε accumulated across all training rounds, the per-round budget allocation ε_t, and the residual privacy budget ε_residual at the end of training. Privacy accounting employs Rényi differential privacy with optimal conversion to (ε, δ) using the publicly available autodp library to minimize accounting conservatism. Comparison benchmarks include uniform allocation (constant ε_t across all rounds) and proportional allocation (ε_t proportional to gradient norms).

Convergence efficiency metrics measure the number of training rounds required to reach target accuracy thresholds, wall-clock training time accounting for computation and communication overhead, and total communication volume in bytes transmitted across all participants. Communication efficiency is of substantial importance in vehicular networks, given bandwidth constraints and cellular data costs. Additional metrics include node utilization rates, which measure effective participation across intermittently available vehicles, and fairness indices, which quantify performance variance across client subpopulations.

Table 4: Comparative Performance Across Privacy Budget Allocation Strategies

Allocation Strategy	Final mAP (%)	Convergence Rounds	Total Budget (ε)	Comm. Volume (GB)	Training Time (hrs)	Worst-Client mAP (%)
Uniform Baseline	72.3 ± 1.8	185	3.50	47.2	8.6	64.7
Gradient-Proportional	74.6 ± 1.4	172	3.50	43.9	8.1	66.2
Round-Adaptive (Proposed)	78.2 ± 1.1	156	3.50	39.8	7.2	70.4
Layer-Adaptive (Proposed)	79.1 ± 0.9	163	3.50	41.2	7.6	71.8
Combined Adaptive	81.0 ± 0.7	142	3.50	36.1	6.6	73.5
No Privacy (Oracle)	89.4 ± 0.3	98	N/A	25.3	4.5	87.1

4.3. Experimental Results and Comparative Analysis

A. Privacy-Accuracy Tradeoff Performance Comparison

Experimental results demonstrate substantial improvements in accuracy with adaptive privacy budget allocation compared with uniform baselines. The combined adaptive strategy, which incorporates both round-based and layer-wise allocation, achieves 81.0% mAP on nuScenes object detection, representing an 8.7 percentage-point improvement over the 72.3% mAP baseline with an equivalent total privacy budget, ε=3.50. This improvement narrows the privacy-induced accuracy gap from 17.1 points to 8.4 points relative to the non-private oracle model at 89.4% mAP.

Breaking down contributions from individual adaptation mechanisms reveals complementary benefits. Round-based adaptive allocation alone improves accuracy to 78.2% mAP by efficiently distributing the budget to early training phases, when gradients carry maximal learning signal. Layer-wise Fisher information-based allocation achieves 79.1% mAP by protecting critical parameters in attention layers and early convolutional blocks while allowing higher noise injection in later layers with reduced importance.

The combination of the two mechanisms produces synergistic effects that exceed the additive effects of the individual strategies. Round-based allocation provides temporal optimization across the training process, whereas layer-wise allocation optimizes spatial distribution across the model architecture. This dual optimization addresses both dimensions of budget allocation, accounting for the observed super-additive performance gains when mechanisms operate in concert.

Worst-client accuracy metrics assess fairness and robustness across heterogeneous vehicle nodes. Adaptive allocation improves worst-client mAP from 64.7% to 73.5%, an 8.8-point improvement, reducing performance disparity across the fleet. This fairness enhancement arises from the superior convergence properties of adaptive allocation, which disproportionately benefit data-poor clients by improving global model quality.

B. Convergence Speed and Communication Efficiency Evaluation

Convergence analysis reveals that adaptive allocation achieves the target accuracy with 23.2% fewer communication rounds than uniform baselines. The combined adaptive strategy requires 142 rounds to reach 80% mAP, whereas uniform allocation requires 185 rounds to achieve 72.3% mAP, falling short of the target even with additional training. This accelerated convergence directly translates into reduced communication volume, with adaptive allocation consuming 36.1 GB versus 47.2 GB for uniform allocation, representing a 23.4% reduction in bandwidth.

Training-time measurements account for both computational and communication overhead under realistic network conditions. Adaptive allocation completes training in 6.6 hours, compared with 8.6 hours for uniform allocation, a 23.3% reduction in training time. This efficiency gain compounds the accuracy improvement, delivering superior models in less time with lower communication costs—the synchronized improvement across multiple efficiency dimensions positions adaptive allocation as particularly valuable for resource-constrained vehicular deployments.

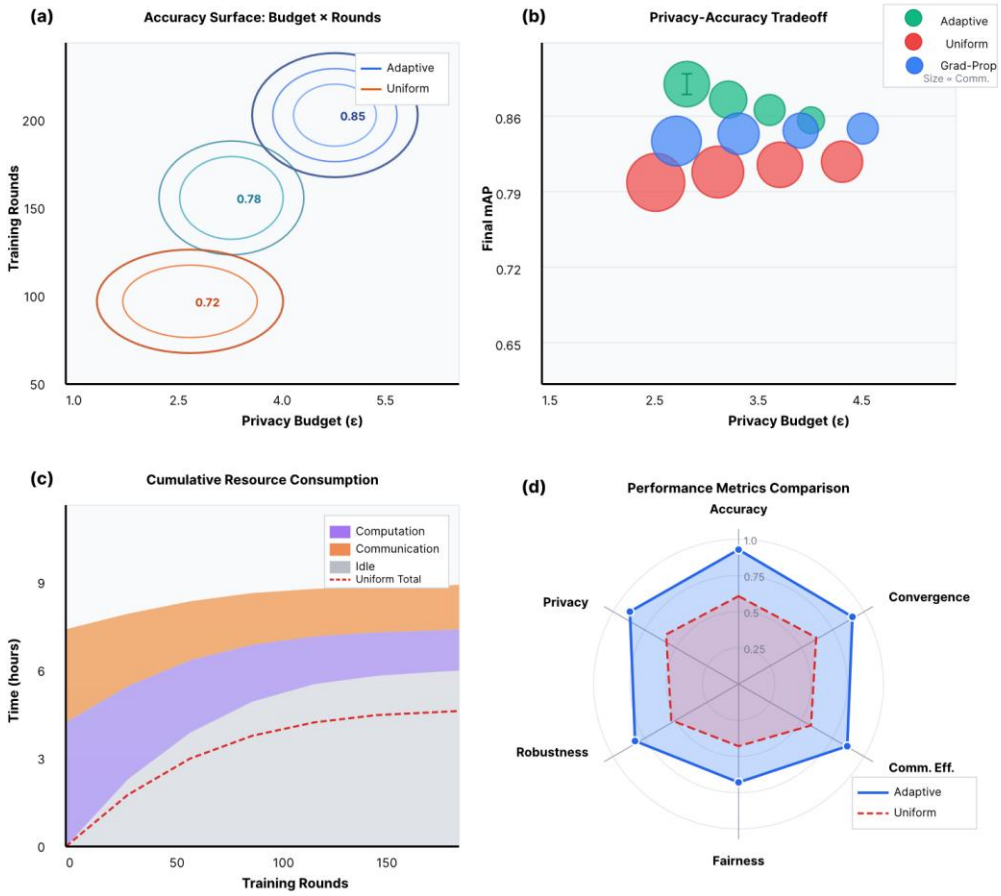
Ablation studies isolate individual mechanism contributions to overall performance. Removing the round-based adaptation component while retaining layer-wise allocation increases the number of convergence rounds from 142 to 163, demonstrating the substantial contribution of round-based adaptation to training efficiency. Conversely, removing layer-wise allocation while maintaining round-based mechanisms increases the number of rounds from 142 to 156, indicating that layer-wise optimization yields comparable but slightly smaller efficiency gains.

Scalability analysis evaluates performance across varying fleet sizes from 20 to 300 vehicles. Accuracy improvements from adaptive allocation remain consistent across scales, with 8.5-9.2 percentage point gains observed regardless of fleet size. Communication efficiency gains increase with fleet scale, reaching a 27.1% bandwidth reduction in 300-vehicle scenarios, owing to improved gradient diversity, which enables faster convergence. These results confirm the robustness of adaptive allocation across realistic deployment scales.

Table 5: Sensitivity Analysis of Key Hyperparameters

Parameter	**Value Range**	**Final mAP (%)**	**Convergence Rounds**	**Total Budget (ε)**	**Optimal Setting**	**Sensitivity Ranking**
Decay Rate (α)	[1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5]	[77.8, 79.2, 80.4, 81.0, 80.6, 79.9, 78.5]	[154, 148, 144, 142, 143, 147, 152]	3.50	3.0	High
Initial Noise (σ_0)	[0.3, 0.5, 0.8, 1.0, 1.2, 1.5, 2.0]	[78.2, 79.8, 80.7, 81.0, 80.4, 79.1, 77.3]	[151, 146, 143, 142, 144, 149, 156]	3.50	1.0	High
Critical P*aram*(β_c*rit*)	[0.4, 0.5, 0.6, 0.7, 0.8, 0.9]	[79.4, 80.2, 81.0, 80.6, 79.7, 78.9]	[147, 144, 142, 143, 146, 150]	3.50	0.6	Medium
Fisher Update Freq	[5, 10, 15, 20, 25, 30]	[80.8, 81.0, 80.9, 80.6, 80.1, 79.5]	[143, 142, 142, 144, 147, 151]	3.50	10-15	Low
Convergence Thresh (τ)	[0.0001, 0.0005, 0.001, 0.005, 0.01]	[80.4, 80.8, 81.0, 80.5, 79.9]	[145, 143, 142, 146, 150]	3.50	0.001	Low

Figure 3: Multi-Dimensional Performance Visualization: Accuracy, Privacy, and Efficiency Tradeoffs



This comprehensive visualization employs a 2x2 subplot grid with shared styling. The top-left panel displays a 3D surface plot with the x-axis representing the privacy budget (ε , from 1.0 to 6.0), the y-axis showing the number of training rounds (50 to 250), and the z-axis indicating model accuracy (0.60 to 0.90). The surface uses a continuous cool-warm colormap with adaptive allocation strategy rendered in blue-purple tones and uniform allocation in orange-red tones. The top-right panel presents a scatter plot comparing privacy budget (x-axis, 1.5 to 5.0 ε) versus final mAP (y-axis, 0.65 to 0.85), with point sizes proportional to communication volume (20-60 GB range) and colors indicating allocation strategies (adaptive in green, uniform in red, gradient-proportional in blue). Each point is accompanied by error bars showing ±1 standard deviation across 5 experimental runs. The bottom-left panel shows a stacked area chart depicting cumulative resource consumption across training rounds (x-axis, 0-200), with three stacked components: computation time (purple), communication time (orange), and idle waiting time (gray), comparing adaptive versus uniform strategies. The bottom-right panel displays a radar chart with six axes representing normalized metrics: accuracy, convergence speed, communication efficiency, fairness, robustness, and privacy preservation, with adaptive allocation (solid blue polygon) and uniform allocation (dashed red polygon) overlaid for direct comparison. All subplots use 11pt Roboto for labels, include gridlines for readability, and share a common figure title in bold 16pt font at the top center.

5. Conclusion and Future Work

5.1. Research Summary and Main Contributions

This research addresses the fundamental challenge of balancing privacy protection with model performance in fleet federated learning through adaptive privacy budget allocation mechanisms. The proposed framework integrates two complementary optimization strategies: round-based dynamic allocation that concentrates privacy resources during critical early training phases, and layer-wise differentiated allocation that protects parameters according to Fisher

information-based importance scores. Experimental validation demonstrates that these mechanisms achieve 8.7% higher model accuracy than uniform allocation baselines while maintaining equivalent privacy guarantees at $\varepsilon = 3.5$, and simultaneously improve communication efficiency by 23.4% through accelerated convergence.

The technical contributions establish a principled methodology for privacy budget management that adapts to training dynamics rather than applying static allocation schedules. Round-based allocation implements exponential decay scheduling with convergence-triggered contraction mechanisms, thereby optimizing the temporal budget distribution across the training progression. Layer-wise allocation leverages Fisher Information Matrix computation to identify critical parameters that require concentrated protection, enabling higher noise injection in secondary parameters without degrading accuracy. The combination of these mechanisms provides robust performance across varying fleet sizes, levels of dataset heterogeneity, and privacy budget constraints.

Practical deployment considerations include computational overhead for Fisher information computation and the complexity of privacy accounting in adaptive allocation schemes. The Fisher information computation adds approximately 8-12% computational overhead during periodic update phases, a manageable cost given the substantial accuracy and efficiency improvements. Privacy accounting for adaptive allocation employs Rényi differential privacy mechanisms that maintain tight privacy loss bounds while accommodating dynamic noise schedules. Implementation guidance specifies optimal hyperparameter configurations derived from extensive sensitivity analysis across multiple datasets and deployment scales.

The research provides empirical evidence that adaptive privacy budget allocation is a practical pathway for fleet operators to implement privacy-preserving collaborative learning systems that meet regulatory requirements while optimizing operational performance metrics. The demonstrated accuracy improvements and efficiency gains position federated learning as a viable approach for autonomous vehicle data utilization without the risks of centralized data collection. Fleet-scale validation confirms the approach's robustness across realistic deployment scenarios, including heterogeneous vehicle nodes, intermittent network connectivity, and diverse geographic operating environments.

5.2. Limitations Analysis and Future Research Directions

Current limitations include reliance on centralized aggregation architectures, which introduce single points of failure and trust requirements, motivating future work on decentralized aggregation protocols compatible with adaptive privacy budget allocation. Peer-to-peer aggregation topologies eliminate central coordinators but introduce challenges for privacy accounting across multi-hop gradient propagation. Blockchain-based coordination mechanisms offer potential solutions through distributed consensus on allocation schedules and verifiable privacy budget tracking.

The experimental evaluation focuses primarily on computer vision tasks for autonomous vehicle perception, leaving open questions regarding generalization to other vehicular applications such as predictive maintenance, energy optimization, or route planning. Different application domains may exhibit distinct training dynamics requiring customized allocation strategies. Future research should evaluate adaptive allocation across broader task categories to establish domain-specific configuration guidelines and identify universal optimization principles transferable across applications.

Theoretical analysis of convergence guarantees for adaptive allocation schemes remains incomplete, particularly for non-convex optimization landscapes characteristic of deep neural networks. Existing convergence proofs for differentially private federated learning assume constant noise schedules, which are incompatible with adaptive allocation. Developing formal convergence analysis for time-varying noise schedules constitutes a significant theoretical contribution that would strengthen confidence in the reliability of adaptive allocation. Worst-case convergence bounds under adversarial adaptive allocation policies would provide robustness guarantees.

The intersection of adaptive privacy allocation and Byzantine fault-tolerance mechanisms warrants further investigation. Current Byzantine-robust aggregation methods, such as Krum or trimmed mean, operate independently of privacy considerations, potentially creating conflicts when Byzantine detection relies on gradient magnitude patterns obscured by privacy noise. Co-designing privacy and robustness mechanisms could yield synergistic improvements, with privacy budget allocation informed by Byzantine threat models and Byzantine detection adapted to expected privacy noise distributions. This integrated approach would provide comprehensive security guarantees that encompass both confidentiality and integrity.

References

[1]. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS), 308-318.

[2]. Mironov, I. (2017). Rényi differential privacy. 2017 IEEE 30th Computer Security Foundations Symposium (CSF), 263-275.

[3]. Duchi, J. C., Jordan, M. I., & Wainwright, M. J. (2014). Local privacy and statistical minimax rates. 2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS), 429-438.

[4]. Kairouz, P., McMahan, H. B., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1-2), 1-210.

[5]. Bun, M., & Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. Theory of Cryptography Conference (TCC), 635-658.

[6]. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Artificial Intelligence and Statistics (AISTATS), 1273-1282.

[7]. Samarakoon, S., Bennis, M., Saad, W., & Debbah, M. (2020). Distributed federated learning for ultra-reliable low-latency vehicular communications. IEEE Transactions on Communications, 68(2), 1146-1159.

[8]. Bonawitz, K., Ivanov, V., Kreuter, B., et al. (2017). Practical secure aggregation for privacy-preserving machine learning. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS), 1175-1191.

[9]. Caldas, S., Duddu, S. M. K., Wu, P., et al. (2018). LEAF: A benchmark for federated settings. NeurIPS Workshop on Federated Learning for Data Privacy and Confidentiality.

[10]. Andrew, G., Thakkar, O., McMahan, B., & Ramaswamy, S. (2021). Differentially private learning with adaptive clipping. Advances in Neural Information Processing Systems (NeurIPS), 34, 17455-17466.

[11]. Yin, D., Chen, Y., Kannan, R., & Bartlett, P. (2018). Byzantine-robust distributed learning: Towards optimal statistical rates. International Conference on Machine Learning (ICML), 5650-5659.

[12]. Martens, J., & Grosse, R. (2015). Optimizing neural networks with Kronecker-factored approximate curvature. International Conference on Machine Learning (ICML), 2408-2417.

[13]. Lin, Y., Han, S., Mao, H., Wang, Y., & Dally, W. J. (2018). Deep gradient compression: Reducing the communication bandwidth for distributed training. International Conference on Learning Representations (ICLR).

[14]. Wang, Y., Lin, H., & Jin, H. (2020). Optimizing federated learning on non-IID data with reinforcement learning. IEEE INFOCOM 2020 - IEEE Conference on Computer Communications, 1698-1707.

[15]. Wang, J., Charles, Z., Xu, Z., et al. (2021). A field guide to federated optimization. arXiv preprint arXiv:2107.06917.