Federated Learning in LEO Constellations: A Simulation-Based Study*

YouJeong Lee[1*], Jakyoung Yun[1*], and Yeon-sup Lim[2†]

[1]Sungshin Women's University, Seoul, South Korea
(20211087, 20221114, ylim)@sungshin.ac.kr
Keywords: LEO Satellites, Federated Learning, Simulation Analysis

Abstract

This study presents a simulation of Federated Learning (FL) in a Low Earth Orbit (LEO) satellite network to analyze how intermittent connectivity and heterogeneous training times affect learning efficiency. Using an event-driven simulator implemented for FL in LEO networks, we evaluate three scenarios representing different levels of connectivity and computational heterogeneity. The results show that intermittent and heterogeneous conditions significantly slow convergence compared to conventional FL settings.

1 Introduction

Low Earth Orbit (LEO) satellite networks have emerged as a next-generation communication infrastructure that enables low-latency, wide-area connectivity. However, directly transmitting sensitive data between satellites and ground stations remains challenging due to limited bandwidth, frequent link interruptions, and privacy concerns. Federated Learning (FL), which keeps data local while exchanging only model parameters, offers a promising solution by enabling collaborative learning without raw data transfer in LEO environments. However, traditional FL procedures are not well suited to real-world LEO settings because of intermittent connectivity, heterogeneous onboard computation, and asynchronous orbital phases [?]. To systematically evaluate FL performance under such conditions, we develop an event-driven simulator that quantifies accuracy, idle ratio, and convergence behavior across various operational scenarios.

2 Implementation

We develop a two-stage LEO satellite-based FL simulator to analyze how intermittent connectivity and heterogeneous local training affect model convergence. In the first stage, the simulator generates an FL event trace based on the connectivity between multiple LEO satellites and ground stations, which is determined by each satellite's orbital position. It discretizes time into one-minute intervals and computes each satellite's position based on orbital dynamics equations. At every time step, the simulator determines whether a satellite is within communication range of any ground station. Using these connectivity results and a given FL configuration, it records detailed events including link establishment and disconnection, global model aggregation, and local model reporting. In the second stage, the simulator performs the actual training process according to the generated event trace. It then tracks and outputs key performance metrics, such as accuracy and loss, over time, enabling quantitative evaluation of how communication interruptions and local heterogeneity influence FL convergence. Figure 1 shows a screenshot of the simulator visualizing connectivity between satellites and the ground station, while Table 1 presents an example of the generated event trace.

3 Experiments

Using the implemented simulator, we investigate how communication availability and computational heterogeneity affect global model convergence. Each satellite trains an identical multilayer perceptron (MLP) with a 100-neuron hidden layer, ReLU activation, and the Adam optimizer (learning rate = 0.001, batch size = 32). Every satellite holds 2,000 non-IID samples from the UNR-IDD network intrusion dataset [?] to represent realistic data heterogeneity. When a satellite establishes a connection with a ground station, it immediately downloads the latest global model, performs one local training epoch on its assigned dataset, and uploads the updated parameters while the link remains stable. Once all participating satellites complete their uploads, the ground station updates the global model through weighted averaging. Because the updated model is distributed only after all uploads are received—and only to satellites currently connected—those

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-44, December 16-18, 2025, Sapporo, Japan.

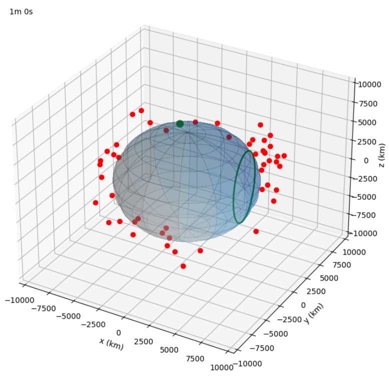

Figure 1: LEO satellite orbits around Earth.

Time (min)	Satellite	Event
6	sat_4	GS connect
6	sat_4	get_global
11	sat_2	GS connect
11	sat_2	get_global
13	sat_6	GS connect
13	sat_6	get_global
17	sat_7	GS connect
17	sat_7	get_global
22	sat_1	GS connect
22	sat_1	get_global
30	sat_2	local
45	sat_4	local

GS connect – connects to Ground Station.
get_global – downloads global model.

local – performs local training.

Table 1: Satellite Event Log

that finish training early may remain idle until synchronization is complete [?]. We evaluate three scenarios reflecting different connectivity patterns and computational capacities. *Always* represents an idealized upper bound with continuous connectivity. *Base* models intermittent contact windows while maintaining identical 10-minute local training durations for all satellites. *Var* introduces heterogeneous local training times ranging from 5 to 15 minutes to emulate varying onboard computation speeds.

4 Results

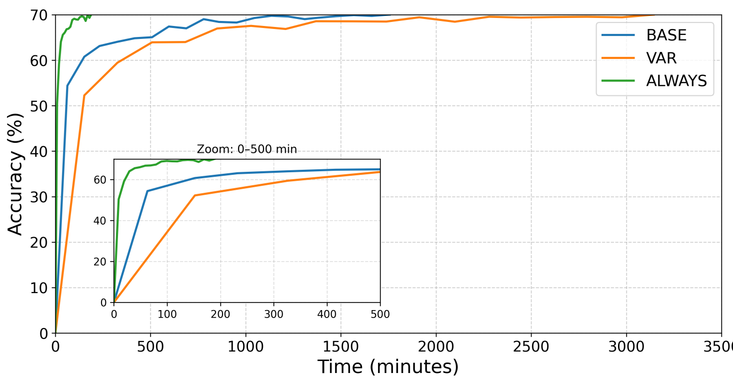

Figure 2: Accuracy over time

In our experiments, we investigate two key evaluation metrics: (1) convergence time, defined as the total simulated duration required to reach 70% accuracy, and (2) idle ratio, defined as the proportion of waiting time to total operation time. These metrics enable a quantitative comparison of synchronization inefficiency across scenarios under consistent model and dataset conditions. The *Always* scenario achieves the target accuracy within approximately 3 hours, demonstrating near-ideal synchronous learning. In contrast, the *Base* scenario requires about 29 hours due to prolonged waiting between contact intervals, while the *Var* scenario shows the slowest convergence at roughly 52 hours, about 1.8× slower than *Base*. The slow convergence in *Var* mainly results from satellites with longer local training times that frequently miss upload opportunities, delaying global aggregation. The average idle ratio increases from nearly 0% in *Always* to 45% in *Base* and 68% in *Var*, confirming that limited communication availability and heterogeneous computation jointly contribute to slower convergence. These results highlight that, even with identical datasets and models, connectivity constraints and resource imbalance critically degrade the learning efficiency of LEO-based FL systems. As future work, we will explore asynchronous model aggregation and adaptive scheduling mechanisms in LEO-based FL systems.