Federated Learning-Based Cross-layer Security Design for Satellite Networks

Zhisheng Yin[1], Yonghong Liu[2], Nan Cheng[1(✉)], Linlin Liang[2], Wenbin Sun[3], and Tom H. Luan[4]

[1] School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
zsyin@xidian.edu.cn, dr.nan.cheng@ieee.org
[2] School of Cyber Engineering, Xidian University, Xi'an 710071, China
23151214131@stu.xidian.edu.cn, llliang@xidian.edu.cn
[3] School of Electronics and Information, Northwestern Polytechnical University, 710072 Xi'an, China
sunwenbin@nwpu.edu.cn
[4] School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China
tom.luan@xjtu.edu.cn

Abstract. The extensive coverage of satellite networks robustly supports federated learning (FL) in multiple domains. This combination protects user privacy and enables extensive data training, with promising applications in remote healthcare, smart agriculture, and environmental monitoring. However, existing FL primarily focuses on data training and aggregation, with less attention given to the secure transmission of model data during upload and download processes. This paper explores cross-layer security in satellite networks, focusing on the physical and application layers. We propose a beamforming optimization scheme based on unsupervised neural network to guarantee secure transmissions without compromising FL training performance. Simulation results underscore the efficacy of our approach in securing physical layer transmissions and affirm its practicality in maintaining robust FL training outcomes.

Keywords: Satellite networks · Federated learning · Cross-layer security · Unsupervised learning

1 Introduction

In recent years, the proliferation of satellite networks (SN) has played a pivotal role in achieving ubiquitous global connectivity, particularly in remote and underserved regions. These networks, encompassing a vast array of connections, harbor substantial data volumes, significantly benefiting machine learning (ML)

H.-H. Chen and W. Meng (Eds.): WiSATS 2024, LNICST 605, pp. 65–76, 2025.
https://doi.org/10.1007/978-3-031-86196-3_6

applications driven by data [1]. However, due to privacy concerns inherent within SN, direct transmission of user data is impractical. Federated learning , as a distributed ML paradigm, emerges as a promising solution to enhance data privacy and reduce the latency associated with centralized data processing [2]. FL involves training across multiple dispersed edge devices or servers that hold local data samples without the need to exchange these samples [3]. This methodology is particularly suited to SN where data privacy and transmission costs are major concerns. Several studies have explored the initial integration of FL with SN, utilizing the distributed nature of these networks to enhance ML models by locally training on satellites and aggregating them through a central server. Notably, the authors of [4] introduced FedSat, which incorporates an asynchronous aggregation algorithm and a corresponding communication protocol, effectively accelerating FL training speeds. Addressing communication challenges in SN, [5] utilized optical relay links to reduce communication latency, alongside effective scheduling strategies for servers and Low Earth Orbit (LEO) satellites to minimize communication costs during FL training. Further, [6] tackled the issue of limited communication bandwidth by proposing SatelliteFL, which accelerates model convergence and enhances bandwidth utilization.

Despite these advancements, most research has primarily focused on optimizing resource allocation, improving data transmission efficiency, and enhancing the robustness of communication links under the unique constraints of SN. However, there remains a significant gap in discussions on the security of model parameter transmission between satellites and ground stations. Transmitting trained models via wireless communications to a central aggregator introduces potential vulnerabilities. Attackers could intercept or eavesdrop on the transmitted model parameters, leading to the leakage of sensitive user information. Some studies have addressed the security aspects of FL, where [7,8] introduced a method combining coding with differential privacy to secure the uplink transmission of model parameters, balancing model accuracy and security. In [9], the authors proposed uncoded wireless FL, implementing differential privacy and adaptive power control for secure transmission. Additionally, [10] focused on the physical layer security of FL in medical data analysis scenarios and proposed a framework based on clustering, along with a security analysis. The aforementioned articles focus on studying the security issues in the transmission process of federated learning models. However, the aforementioned works either focus primarily on physical layer security or on federated learning, without highlighting the importance and interrelationship of studying both aspects together.

Addressing these challenges, this paper proposes a novel secure transmission scheme for FL models in SN. Based on the channel state information (CSI) of both legitimate and wiretap channels, our approach optimizes the beamforming vectors of each client to maximize the minimum secrecy rate, thereby achieving physical layer secure transmission. This method offers superior physical layer security performance compared to traditional approaches. Furthermore, our results, which consider the actual transmission effects on the received model parameters, demonstrate that while there is a slight sacrifice in FL learning

performance, the method significantly reduces the risk of eavesdropping. The main contributions of this paper are outlined as follows:

- **Novel Secure Transmission Scheme for FL in SN:** We introduce a new secure transmission scheme specifically designed for FL models in SN. This scheme leverages the CSI of both legitimate and wiretap channels. By optimizing the beamforming vectors of each client with the objective of maximizing the minimum secrecy rate, the scheme ensures enhanced physical layer security during the transmission of FL models.
- **Enhanced Physical Layer Security Performance:** The proposed method outperforms traditional physical layer security approaches by providing superior security performance. This is achieved through the strategic optimization of beamforming vectors based on the CSI, which not only secures the data transmission against eavesdropping but also maximizes the efficiency of the transmission by focusing the signal power in the direction of legitimate receivers while minimizing leakage to potential eavesdroppers.
- **Comprehensive Experimental Evaluation:** We conducted rigorous experiments to evaluate the impact of the secure transmission scheme on the learning performance and security of FL in satellite networks. The results indicate that our proposed unsupervised learning method outperforms traditional physical layer security methods in terms of secrecy rate performance, and the corresponding FL training performance is closer to the ideal scenario. This demonstrates the effectiveness and feasibility of our proposed method.

In Sect. 2, we introduce the system model, including the federated learning model, signal modeling, and problem formulation. In Sect. 3, we propose an unsupervised learning method for optimizing beamforming vectors and explain the integration of physical layer secure transmission schemes with federated learning training. Section 4 presents the simulation results and performance evaluations, followed by the conclusion and future research directions in Sect. 5.

2 System Model and Problem Formulation

2.1 FL Model

In this paper, we consider a scenario where each ground Access Point (AP) functions as a client in a FL framework. To achieve the aggregation of neural network parameters trained by each client, we utilize a LEO satellite as the FL server. The selection of LEO is due to its extensive coverage and ability to simultaneously receive neural networks transmitted by multiple ground APs and aggregate them effectively.

FL aims to enable collaborative model training across multiple decentralized devices without sharing raw data. Each client performs local training on its dataset, and only the model parameters are communicated to the server. The process begins with initializing the global model at the LEO server, which is then distributed to all ground APs. Each AP, acting as a client, trains the model locally using its dataset. Once local training is completed, the APs upload their

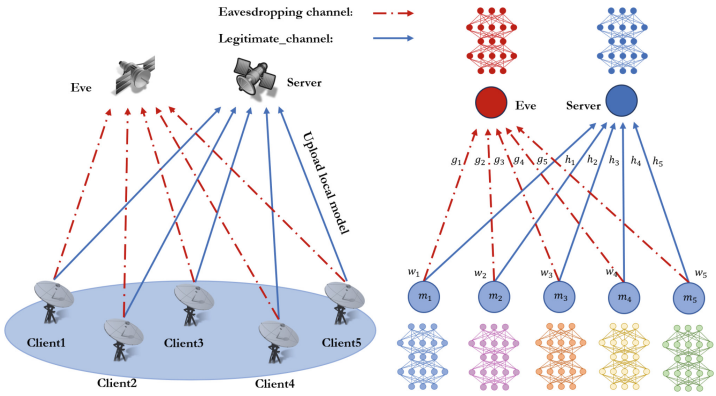

Fig. 1. Federated Learning with Satellite Networks

local models to the LEO satellite simultaneously. The LEO server aggregates these local models and sends the updated global model back to the ground APs. This process is iterated until the FL training converges, resulting in a robust and generalized global model. For the aggregation of the model parameters at the LEO satellite, we employ the Federated Averaging (FedAvg) algorithm. The principle of FedAvg is to compute a weighted average of the model parameters from all clients, effectively integrating the contributions from each client based on the size of their local datasets. The global loss function $F(\mathbf{m}^g)$ is defined as follows:

$$F(\mathbf{m}^g) = \frac{1}{S}\sum_{k=1}^{K}|S_k|F_k(\mathbf{m}^\ell) \tag{1}$$

where, $\mathbf{m}^g$ represents the global model parameter, and $S = \sum_k |S_k|$ refers to the total size of the distributed datasets. The local loss function for client k, $F_k(\mathbf{m})$, is defined as:

$$F_k(\mathbf{m}) = \frac{1}{|S_k|}\sum_{(\mathbf{u}_{k,j},\mathbf{v}_{k,j})\in S_k} f(\mathbf{m};\mathbf{u}_{k,j},\mathbf{v}_{k,j}) + \lambda R(\mathbf{m}) \tag{2}$$

In this context, $f(\mathbf{m};\mathbf{u}_{k,j},\mathbf{v}_{k,j})$ is the sample-wise loss function, $R(\mathbf{m})$ is a strongly convex regularization function, and $\lambda \geq 0$ is a regularization parameter used to prevent model overfitting during training. The goal of FL is to minimize the global loss function, which is given by:

$$\mathbf{m}^\star = \arg\min_{\mathbf{m}^g} F(\mathbf{m}^g) \tag{3}$$

The local model obtained by any client k in an iteration round μ is denoted as $\mathbf{m}^\ell_{k,\mu}$, and the aggregated global model $\mathbf{m}^g_\mu$ is computed as:

$$\mathbf{m}^g_\mu = \sum_{k=1}^{K}\mathbf{m}^\ell_{k,\mu}\frac{|S_k|}{S} \tag{4}$$

To minimize $F(\mathbf{m}^g)$, gradient descent is employed to update local model parameters, which are then utilized to update the global model parameters via Eq. (4). The local model parameters update process is described as follows:

$$\mathbf{m}_{k,\mu}^{\ell} = \mathbf{m}_{\mu-1}^{g} - \frac{\lambda_{lr}}{|S_k|} \sum_{j=1}^{|S_k|} \nabla f(\mathbf{m}_{\mu-1}^{g}; \mathbf{u}_{k,j}, \mathbf{v}_{k,j}) \tag{5}$$

where λ_{lr} denotes the learning rate, and $\nabla f(\mathbf{m}_{\mu-1}^{g}; \mathbf{u}_{k,j}, \mathbf{v}_{k,j})$ is the gradient of the loss function $f(\mathbf{m}_{\mu-1}^{g}; \mathbf{u}_{k,j}, \mathbf{v}_{k,j})$ with respect to $\mathbf{m}_{\mu-1}^{g}$. By repeating this process of local training, uploading, and aggregation, the FL training converges, resulting in an optimized global model that leverages the diverse datasets distributed across multiple ground APs. This approach ensures a scalable and efficient training process, facilitated by the comprehensive coverage of the LEO satellite.

2.2 Signal Model

In the context of Space-Air-Ground integrated networks, we investigate federated learning models in satellite-ground networks, as illustrated in Fig. 1. In low Earth orbit satellites, there are also eavesdropping satellites that exploit the broadcast nature of wireless transmissions to intercept information. Therefore, we focus on the security of the uplink transmission of local model parameters during the federated learning training process.

We assume that all clients are on the same frequency band, and the received signal at the satellite server can be represented as:

$$y = \sum_{k \in K} \mathbf{h}_k^{\dagger} \mathbf{w}_k x_k + n \tag{6}$$

where $\mathbf{h}_k$ denotes the CSI from the k^{th} client to satellite, $(\cdot)^{\dagger}$ denotes the Hermitian transpose, $\mathbf{w}_k \in \mathbb{C}^{N \times 1}$ represents the beamforming vector at the k^{th} client for shaping the uploading signal, x_k contains the confident information which represents the trained local model parameters, and n denotes the noise received by satellite.

Due to the openness of the wireless channel, the eavesdropping signal received by the non-cooperative satellite in the system can be represented as:

$$y_e = \sum_{k \in K} \mathbf{g}_k^{\dagger} \mathbf{w}_k x_k + n_e \tag{7}$$

where $\mathbf{g}_k$ denotes the CSI from the k^{th} client to Eve and n_e represents the noise received by Eve.

2.3 Problem Formulation

Based on (6) and (7), we can calculate the uplink SINRs of the clients at satellite server and eavesdropping satellite, which are obtained as:

$$\gamma_k = \frac{\left\|\mathbf{h}_k^\dagger \mathbf{w}_k\right\|^2}{\sum\limits_{i\neq k, k\in K}\left\|\mathbf{h}_i^\dagger \mathbf{w}_i\right\|^2 + \delta_k^2} \tag{8}$$

$$\gamma_{k,e} = \frac{\left\|\mathbf{g}_k^\dagger \mathbf{w}_k\right\|^2}{\sum\limits_{i\neq k, k\in K}\left\|\mathbf{g}_i^\dagger \mathbf{w}_i\right\|^2 + \delta_e^2} \tag{9}$$

where $\delta_k^2 = \delta_e^2 = 1$ represent the noise power. This simplification assumes equal noise power for both variables.

The secrecy rate of transmission can be obtained as:

$$R_k = \left[\log_2\left(1+\gamma_k\right) - \log_2\left(1+\gamma_{k,e}\right)\right]^+. \tag{10}$$

To enhance the secrecy rate of the uplink from clients to satellite server and ensure fairness in confidentiality, we design a problem formulation aimed at maximizing the minimum secrecy rate of the uplink transmission, which can be mathematically expressed as:

$$\mathcal{P}1: \quad \mathrm{MaxMin}_{\{\mathbf{w}_k\}}\{R_k\}, \tag{11}$$

$$\text{s.t.}: \quad \|\mathbf{w}_k\|^2 \leq P, \tag{11a}$$

$$\mathbf{w}_k \succ \mathbf{0}, \tag{11b}$$

in which (11a) constrains the total power of the system using P, and (11b) constrains the beamforming at the clients.

3 Federated Learning-Based Cross-Layer Security Design

In this section, we propose a beamforming vector optimization method to solve the non-convex problem presented in (11) and explain how to associate the physical layer secure transmission scheme with federated learning training performance.

3.1 Unsupervised Learning Methods for Beamforming

In this study, we utilize a deep complex network for training. We employ unsupervised deep learning to optimize the beamforming vectors in a wireless communication system. To address optimization problem $\mathcal{P}1$, the neural network needs to be trained based on legitimate and wiretap channel parameters. Since both legitimate and eavesdropping channel parameters are complex matrices, we utilize a deep complex network for training to avoid loss of important information due to operations such as extraction, concatenation, and unfolding.

The input to the network is a complex matrix $\mathrm{H} = \mathrm{h}_{\mathrm{real}} + i\mathrm{h}_{\mathrm{imag}}$, where $\mathrm{h}_{\mathrm{real}}$ and $\mathrm{h}_{\mathrm{imag}}$ represent the real and imaginary parts of the legitimate and wiretap channel parameters, respectively. For convolution operations in the complex domain, the complex vector $\mathrm{z} = \mathrm{x} + \mathrm{iy}$ is processed as:

$$\mathrm{H} * \mathrm{z} = (\mathrm{h}_{\mathrm{real}} * \mathrm{x} - \mathrm{h}_{\mathrm{imag}} * \mathrm{y}) + \mathrm{i}(\mathrm{h}_{\mathrm{imag}} * \mathrm{x} + \mathrm{h}_{\mathrm{real}} * \mathrm{y}) \tag{12}$$

Complex convolution layers can thus perform equivalent operations by leveraging traditional real-valued convolution layers. We set up four complex convolution layers for training, with the output channel numbers of the first, second, and third layers being 32, 16, and 8, respectively. Each layer uses the LeakyReLU activation function, defined as:

$$\mathrm{LeakyReLU}(\mathrm{x}) = \begin{cases} \mathrm{x}, & \text{if } \mathrm{x} > 0 \\ \alpha\mathrm{x}, & \text{if } \mathrm{x} \leq 0 \end{cases} \tag{13}$$

Compared to the ReLU function, LeakyReLU effectively addresses the issue of gradient vanishing caused by negative input values. The Adam algorithm is employed throughout the training process to fine-tune the neural network's parameters and weights, ensuring stable convergence of the objective function and optimal secrecy rate performance. The network outputs beamforming vectors for five users. The neural network parameters, θ, are updated according to the gradient update rule:

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{14}$$

where, α represents the learning rate, ϵ is a small constant to prevent division by zero, and $\hat{m}_t$ and $\hat{v}_t$ are bias corrected estimates of the first moment (mean) and the second moment (uncentered variance), defined as:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{15}$$

where β_1 and β_2 represent the decay rates for the first and second moment estimates, respectively. The estimates m_t and v_t are defined as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\mathcal{L}(\boldsymbol{\theta}_t) \tag{16}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\mathcal{L}(\boldsymbol{\theta}_t)^2 \tag{17}$$

Since there is a constraint equation in the optimization problem $\mathcal{P}1$, a straightforward method is to use the Lagrangian dual method to transform the constrained optimization problem into an unconstrained one. The neural network training optimizes the Lagrangian dual function. However, the non-interpretability of the NN based on the statistical learning principle implies that the NN can only probabilistically optimize the Lagrangian dual function and cannot ensure that the output will always satisfy the constraint. Therefore, we

design an efficient activation function $\mathcal{S}$ to ensure that the NN output adheres to the constraint.

The activation function $\mathcal{S}$ processes a complex vector as input and outputs a complex vector. If the square of the two-norm of the complex vector is less than P, it directly outputs the input complex vector. If the square of the two-norm exceeds P, it scales the complex vector by dividing it by the square of the two-norm and multiplying by P. This ensures that the square of the two-norm of the output complex vector is always less than or equal to P. Formally, $\mathcal{S}$ is defined as:

$$\mathcal{S}(\mathbf{w}_k) = \begin{cases} \mathbf{w}_k, & \text{if } \|\mathbf{w}_k\|^2 \leq P \\ \mathbf{w}_k \frac{\sqrt{P}}{\|\mathbf{w}_k\|}, & \text{if } \|\mathbf{w}_k\|^2 > P \end{cases} \tag{18}$$

The derivative of $\mathcal{S}$ with respect to the input complex vector $\mathbf{w}_k$ is given by:

$$\frac{\partial \mathcal{S}(\mathbf{w}_k)}{\partial \mathbf{w}_k} = \begin{cases} 1, & \text{if } \|\mathbf{w}_k\|^2 \leq P \\ \frac{\sqrt{P}}{\|\mathbf{w}_k\|} - \frac{\mathbf{w}_k(\mathbf{w}_k^H \mathbf{w}_k)}{\|\mathbf{w}_k\|^3}, & \text{if } \|\mathbf{w}_k\|^2 > P \end{cases} \tag{19}$$

This piecewise derivative confirms that $\mathcal{S}$ is differentiable. By incorporating the activation function $\mathcal{S}$ and defining the loss function as $\mathcal{L}(\boldsymbol{\theta}_t) = -\min\ R_k$, we ensure that the NN training process adheres to the constraints and achieves optimal performance. This allows for chain derivation to obtain the derivative of the NN parameter with respect to the loss function, facilitating unsupervised NN training.

To ensure the magnitude constraint on the beamforming vector of each client, we define a mask as follows:

$$\text{mask}_k = \max\left(1, \frac{\|\mathbf{w}_k\|^2}{P}\right) \tag{20}$$

The beamforming vector is then adjusted using the mask

$$\mathbf{w_k}' = \frac{\mathbf{w_k}}{\text{mask}_k} \tag{21}$$

By integrating the activation function $\mathcal{S}$ and the loss function $\mathcal{L}$, we ensure that the NN training process meets the constraints and achieves optimal performance in beamforming vector optimization, thus similar to [11,12] the gradient of objective function with respect to the parameters of NN can be obtained by chain rule, resulting in an unsupervised learning method.

3.2 Combination with Federated Learning

In an ideal federated learning model, the uplink transmission links where each client uploads local model parameters to the server are assumed to be perfectly transmitted. Thus, aggregation at the server can be directly computed using the formula. However, in actual transmission processes, signals are subject to attenuation, noise, and other influences. To simplify this process, it is understandable

to directly add noise to the model parameters based on the SINR calculated from formulas (8) and (9), equivalent to the noise impact received during the uplink transmission process.

However, adding noise to model parameters $\mathbf{m}$ at the data level based on the signal level metric SINR is unreasonable. Therefore, we further model the relationship between SINR and the model parameters through bit error rate (BER) and parameter quantization with bit flipping, which can be calculated using the following formula:

$$\mathrm{BER} = Q\left(\sqrt{2 \cdot \mathrm{SINR}}\right) \tag{22}$$

The Q function is the Gaussian Q-function. Next, we use this BER model for parameter quantization and bit flipping. First, the model parameters are quantized. Assuming the model parameters are represented as a floating-point array $\mathbf{m}$, we quantize them into binary representation:

$$\mathbf{m}_{\mathrm{binary}} = \mathrm{Binary}(\mathbf{m}) \tag{23}$$

where $\mathrm{Binary}(\cdot)$ denotes converting a floating-point number into its binary representation.

According to the BER calculated in (22) and the binary representation of the model parameters generated in (23), we generate a random flipping mask for each bit, defined as:

$$\mathbf{m}_{\mathrm{mask}} = \mathrm{FlipMask}(\mathbf{m}_{\mathrm{binary}}, \mathrm{BER}) \tag{24}$$

The rule for bit flipping the binary model parameters $\mathbf{m}_{\mathrm{binary}}$ according to $\mathbf{m}_{\mathrm{mask}}$ is defined as follows:

$$\mathbf{m}_{\mathrm{binary}}[i] = \begin{cases} 1 - \mathbf{m}_{\mathrm{binary}}[i], & \text{if } \mathbf{m}_{\mathrm{mask}}[i] = 1 \\ \mathbf{m}_{\mathrm{binary}}[i], & \text{otherwise} \end{cases} \tag{25}$$

After the binary model parameters have undergone bit flipping, they can be converted back to floating-point format for model aggregation at the server.

4 Numerical Results

In this section, we evaluate the secrecy performance of clients' uplink transmissions and compare the performance of federated learning models under different conditions. The simulation parameters are configured as follows: five clients are randomly distributed within a 1000 km radius area centered on the satellite server, with the satellite positioned at coordinates (0 km, 0 km, 600 km). All clients are equipped with 4 antennas. The channel power gains from clients to the satellite server and to the eavesdropping satellite at a reference distance of 1 meter are set to -40 dB. The Rician factors for the channels from clients to the satellite server and to the eavesdropping satellite are 10 dB and 5 dB, respectively, for the uplink transmissions.

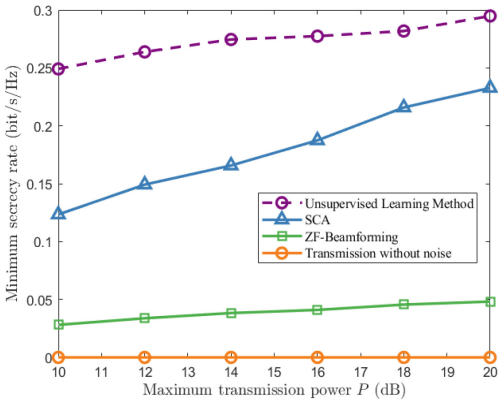

Fig. 2. Maximum transmission power P Vs. the minimum secrecy rate

Figure 2 shows the influence of the maximum transmission power on the minimum secrecy rate. As observed, the secrecy rate performance increases with the maximum transmission power. According to Theorem 1 in [13], as the maximum transmission power increases, more power will be allocated to clients, thereby improving the performance of secrecy rate. Compared to successive convex approximation(SCA) and Zero Forcing(ZF) beamforming, our proposed unsupervised learning method can optimize and achieve beamforming vector schemes with better secrecy rate performance. When PLS is not considered, the signal secrecy rate is essentially zero, making it vulnerable to eavesdropping attack.

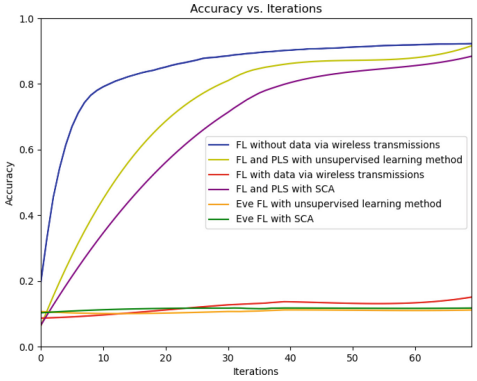

Fig. 3. Comparison of the accuracy

In Fig. 3, it can be observed that the unsupervised learning method we proposed achieves accuracy closer to the ideal scenario than SCA. This demonstrates

that our proposed method can meet physical layer security standards while minimizing the impact on federated learning training accuracy performance. When noise during transmission is considered, federated learning training fails to converge. Similarly, eavesdroppers cannot train based on intercepted signals, demonstrating the effectiveness of the physical layer security transmission scheme in mitigating eavesdropping risks.

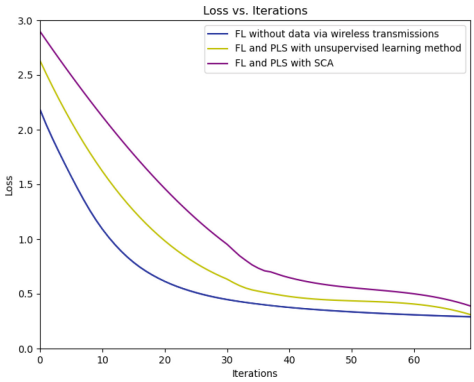

Fig. 4. Comparison of the loss

In Fig. 4, our proposed unsupervised learning method for optimizing beamforming vectors achieves physical layer security transmission while ensuring that the federated learning model achieves loss close to ideal conditions, which outperforms SCA in FL loss performance. Without any processing but considering noise during actual transmission, the FL losses corresponding to both beamforming vector optimization methods at the eavesdropping satellite are too large to be depicted in the figure, indicating their failure to achieve training effectiveness. This underscores the practical significance of beamforming vector optimization for secure model parameter transmission.

5 Conclusion

To achieve cross-layer security in satellite networks, this paper ensures the security of federated data uploads through secure transmission design while also considering FL learning performance. Particularly, an unsupervised learning approach for beamforming optimization in satellite uplinks is proposed, which achieves a superior secrecy rate. Simultaneously, its federated learning model performance decreases only slightly relative to ideal conditions. This demonstrates the feasibility and effectiveness of our proposed method, indicating that it is possible to enhance both secure transmission and federated learning performance in real wireless communication scenarios. In future work, the secure transmission of global model parameters in federated learning will be further investigated.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (No. 62201432, 62071356, and 62101429).

References

1. Cheng, N., et al.: 6g service-oriented space-air-ground integrated network: a survey. Chin. J. Aeronaut. **35**, 1–18 (2022)
2. McMahan, B., Moore, E., Ramage, D., Hampson, S., Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Singh, A., Zhu, J. (eds.) Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 54, pp. 1273–1282. PMLR, 20–22 April 2017
3. Shen, J., et al.: RingSFL: an adaptive split federated learning towards taming client heterogeneity. IEEE Trans. Mob. Comput. **23**(5), 5462–5478 (2024)
4. Razmi, N., Matthiesen, B., Dekorsy, A., Popovski, P.: Ground-assisted federated learning in LEO satellite constellations. IEEE Wirel. Commun. Lett. **11**(4), 717–721 (2022)
5. Chen, C.Y., Shen, L.H., Feng, K.T., Yang, L.L., Wu, J.M.: Edge selection and clustering for federated learning in optical inter-LEO satellite constellation. In: 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1–6 (2023)
6. Yang, C., et al.: Communication-efficient satellite-ground federated learning through progressive weight quantization. IEEE Trans. Mob. Comput. **23**, 1–14 (2024)
7. Zhang, H., Yang, C., Dai, B.: When wireless federated learning meets physical layer security: the fundamental limits. In: IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 1–6 (2022)
8. Zhang, H., Yang, C., Dai, B.: A finite blocklength approach for wireless hierarchical federated learning in the presence of physical layer security. In: IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 1–6 (2023)
9. Liu, D., Simeone, O.: Privacy for free: wireless federated learning via uncoded transmission with adaptive power control. IEEE J. Sel. Areas Commun. **39**(1), 170–185 (2021)
10. Ahmed, J., Nguyen, T.N., Ali, B., Javed, M.A., Mirza, J.: On the physical layer security of federated learning based IoMT networks. IEEE J. Biomed. Health Inform. **27**(2), 691–697 (2023)
11. Wang, X., et al.: Scalable resource management for dynamic MEC: an unsupervised link-output graph neural network approach. In: 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1–6. IEEE (2023)
12. Wang, X., et al.: Joint flying relay location and routing optimization for 6g UAV-IoT networks: a graph neural network-based approach. Rem. Sens. **14**(17), 4377 (2022)
13. Yin, Z., et al.: UAV-assisted secure uplink communications in satellite-supported IoT: secrecy fairness approach. IEEE Internet Things J. **11**(4), 6904–6915 (2024)