Channel Estimation in RIS-Aided Heterogeneous Wireless Networks via Federated Learning

Muhammad Asaad Cheema, *Graduate Student Member, IEEE*, Apoorva Chawla, *Member, IEEE*, Vinay Chakravarthi Gogineni, *Senior Member, IEEE*, and Pierluigi Salvo Rossi, *Senior Member, IEEE*

***Abstract*— Downlink channel estimation in reconfigurable intelligent surface (RIS)-assisted communication systems employing federated learning (FL) is challenging due to communication/computational overhead, users heterogeneity, and vulnerability to malicious users. This letter proposes a novel methodology integrating principal component analysis (PCA)-based clustering with FL, tailored for heterogeneous users. The approach effectively identifies regions and users within the cell while minimizing communication/computational overhead associated with clustering, resulting in accurate, resource-efficient, and secure channel estimation. Simulation results demonstrate that the proposed FL strategy achieves estimation performance comparable to the conventional methods while significantly reducing the communication overhead, enhancing the system security, and handling heterogeneous users.**

***Index Terms*— Channel estimation, federated learning, heterogeneous users, reconfigurable intelligent surfaces.**

I. Introduction

RECONFIGURABLE intelligent surfaces (RISs), thin planar surfaces composed of passive low-cost components, are a pivotal innovation in 6G wireless communications [1], [2]. Each RIS component independently directs the incident signal towards the desired direction by imposing a predefined phase shift, thus enhancing communication quality and performance [3], [4]. A crucial task in RIS-assisted communication systems is to acquire accurate channel state information (CSI), necessary for effective beamforming strategies [5], [6]. Although traditional methods, i.e. least squares (LS) [7] or minimum mean squared error (MMSE) [8], can be employed for cascaded channel estimation, they require significantly higher pilot overhead in RIS-assisted systems than in conventional massive multiple-input multiple-output (MIMO) systems [9]. Codebook-based solutions can reduce the pilot overhead, but need effective codebook generation [10].

Machine learning (ML)-based techniques have been shown to provide effective cascaded channel estimation with reduced pilot overhead [9]. Specifically, these methods employ neural networks to map low-dimensional signals to high-dimensional channels, offering a promising alternative to traditional approaches. Several works [9], [11], [12] exploited deep learning (DL) for efficient estimation of cascaded channels, but always focusing on centralized learning (CL) approaches, requiring all channel data to be collected at the base station (BS) for model training, which leads to significant communication overhead.

Federated learning (FL) can mitigate communication overhead issues by learning a global model without users transmitting their data to the central server [13], [14]. However, simplistic assumptions (e.g., all users/models are trustworthy, all users possess equal computational resources) have limited practical implementations of FL in real-world settings due to suboptimal performance or potential disruptions in presence of heterogeneous or malicious users [15].

We propose a framework for downlink channel estimation in RIS-enabled communications leveraging unsupervised clustering and robust FL to reduce communication overhead, integrate heterogeneous users, and enhance security during the training process. More specifically, the framework includes: (i) communication-efficient one-shot clustering technique based on principal component analysis (PCA); (ii) selective global-model training based on region identification; (iii) model validation checks (MVCs) in each FL iteration to prevent model poisoning. It is worth noticing that we exploit cluster information for building efficiently a single global model. Numerical results from simulations are provided to validate the effectiveness of the proposed framework.

Notation: Bold uppercase/lowercase letters denote matrices/vectors, respectively. $\mathbf{A}_{ij}$ is the (i, j)th entry of the matrix $\mathbf{A}$. The symbols $(\cdot)^T$ and $\otimes$ represent transpose and Kronecker product. $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix with the elements of vector $\mathbf{a}$ on the main diagonal. Calligraphic letters represent sets, except for $\mathcal{L}$ denoting a loss function. $|\mathcal{C}|$ is the cardinality of the set $\mathcal{C}$. $\nabla f(\cdot)$ denotes the gradient of a function.

II. System Model

We consider a narrowband single-cell wireless system with K single-antenna users, one BS equipped with M-antenna uniform planar array (UPA), and a RIS with N elements and dedicated wireless link to the BS. The cell has R regions, each with U_r users ($\sum_{r=1}^{R} U_r = K$) as in Fig. 1. Different regions exhibit different channel characteristics influenced by the angles of arrival/departure of the channels, the distances

Received 23 January 2025; accepted 6 February 2025. Date of publication 13 February 2025; date of current version 11 April 2025. This work was partially supported by the Research Council of Norway under the project ML4ITS within the IKTPLUSS framework. The associate editor coordinating the review of this letter and approving it for publication was W. Liu. *(Corresponding author: Muhammad Asaad Cheema.)*

Muhammad Asaad Cheema, Apoorva Chawla, and Pierluigi Salvo Rossi are with the Department of Electronic Systems, Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: asaad.cheema@ntnu.no; apoorva.chawla@ntnu.no; salvorossi@ieee.org).

Vinay Chakravarthi Gogineni is with the Maersk Mc-Kinney Moller Institute, SDU Applied AI and Data Science, 5230 Odense, Denmark (e-mail: vigo@mmmi.sdu.dk).

Digital Object Identifier 10.1109/LCOMM.2025.3541873

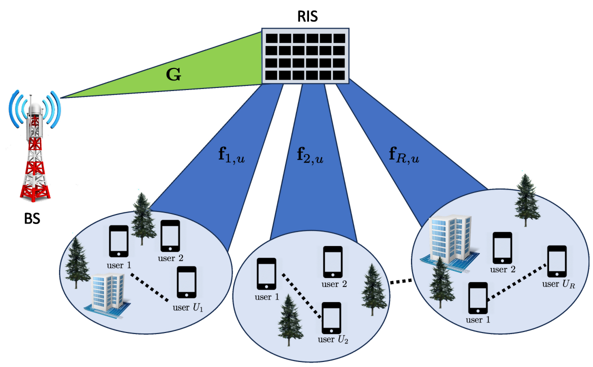


Fig. 1. RIS-aided wireless network.

between BS-user pairs, etc. Mobile users experience significant changes in channel properties when moving from one region to another.

RIS-aided systems operating in frequency division duplex (FDD) mode [16], [17] experience non-reciprocal downlink and uplink channels. We focus on acquiring the CSI of the reflecting link.[1] Assuming that the BS interacts with the users via RIS while ignoring the direct channel, the received signal $y_{r,m}$ at the mth user in the rth region is

$$y_{r,m} = \mathbf{f}_{r,m}^T \mathbf{\Phi} \mathbf{G} \mathbf{w}_{r,m} x_{r,m} + n_{r,m}, \tag{1}$$

where $x_{r,m} \in \mathbb{C}$ is the transmitted signal, $\mathbf{w}_{r,m} \in \mathbb{C}^{M\times 1}$ is the precoding vector at the BS, $\mathbf{G} \in \mathbb{C}^{N\times M}$ is the channel matrix between the RIS and the BS, $\mathbf{f}_{r,m} \in \mathbb{C}^{N\times 1}$ is the channel vector from the RIS to the mth user in the rth region, $n_{r,m} \sim \mathcal{CN}(0, \sigma_n^2)$ is the additive white Gaussian noise (AWGN). Also, $\mathbf{\Phi} = \mathrm{diag}(\phi_1, \phi_2, \ldots, \phi_N) \in \mathbb{C}^{N\times N}$ is the reflecting matrix at the RIS, where ϕ_n is the reflection coefficient of the nth element. According to the Saleh-Valenzuela model,

$$\mathbf{G} = \sqrt{\frac{MN}{L_G}} \sum_{\ell_1=1}^{L_G} \alpha_{\ell_1}^G \mathbf{a}\left(\vartheta_{\ell_1}^{G_r}, \psi_{\ell_1}^{G_r}\right) \mathbf{b}\left(\vartheta_{\ell_1}^{G_t}, \psi_{\ell_1}^{G_t}\right)^T, \tag{2}$$

where $\alpha_{\ell_1}^G$ is the complex gain, L_G is the number of paths between the BS and the RIS, $\vartheta_{\ell_1}^{G_r}$ (resp. $\psi_{\ell_1}^{G_r}$) and $\vartheta_{\ell_1}^{G_t}$ (resp. $\psi_{\ell_1}^{G_t}$) are the azimuth (resp. elevation) angles for the ℓ_1th path at the RIS and the BS. Analogously

$$\mathbf{f}_{r,m} = \sqrt{\frac{N}{L_{r,m}}} \sum_{\ell_2=1}^{L_{r,m}} \alpha_{\ell_2}^{r,m} \mathbf{a}\left(\vartheta_{\ell_2}^{r,m}, \psi_{\ell_2}^{r,m}\right), \tag{3}$$

where $\alpha_{\ell_2}^{r,m}$, $\vartheta_{\ell_2}^{r,m}$, and $\psi_{\ell_2}^{r,m}$ are the complex gain, the azimuth angle, and the elevation angle for the ℓ_2 path at the RIS, $L_{r,m}$ is the number of paths between the RIS and the mth user in the rth region, $\mathbf{a}(\vartheta, \psi) \in \mathbb{C}^{N\times 1}$ and $\mathbf{b}(\vartheta, \psi) \in \mathbb{C}^{M\times 1}$ are the normalized steering vectors for the RIS and the BS. In case of an $N_1 \times N_2$ UPA with half-wavelength spacing [18]

$$\mathbf{a}(\vartheta, \psi) = \frac{1}{\sqrt{N_1 N_2}} \left[e^{-j\pi \cos(\psi)\mathbf{n}_1}\right] \otimes \left[e^{-j\pi \sin(\psi)\cos(\vartheta)\mathbf{n}_2}\right], \tag{4}$$

where $\mathbf{n}_i = [0, 1, \ldots, N_i - 1]^T$. Eq. (1) can be expressed as

$$y_{r,m} = \boldsymbol{\phi}^T \mathbf{H}_{r,m} \mathbf{w}_{r,m} x_{r,m} + n_{r,m}, \tag{5}$$

where $\mathbf{H}_{r,m} = \mathrm{diag}(\mathbf{f}_{r,m}^T)\mathbf{G} \in \mathbb{C}^{N\times M}$ is the downlink cascaded channel matrix and $\boldsymbol{\phi} = [\phi_1, \phi_2, \ldots, \phi_N]^T \in \mathbb{C}^{N\times 1}$.

The BS transmits pilots (without loss of generality set to 1) to users using Q time slots, with the received signal being

$$\mathbf{y}_{r,m}^p = \mathbf{\Theta} \mathbf{H}_{r,m} \mathbf{w}_{r,m} + \mathbf{n}_{r,m}, \tag{6}$$

where $\mathbf{y}_{r,m}^{p,T} = [y_{r,m,1}^p, y_{r,m,2}^p, \ldots, y_{r,m,Q}^p]$ collects the received signals, $\mathbf{\Theta}^T = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_Q]$ collects the RIS coefficients, and $\mathbf{n}_{r,m}^T = [n_{r,m,1}, n_{r,m,2}, \ldots, n_{r,m,Q}]$ collects the noise. Denoting $\mathbf{h}_{r,m} = \mathrm{vec}(\mathbf{H}_{r,m}) \in \mathbb{C}^{MN\times 1}$ and $\mathbf{\Psi}_{r,m} = (\mathbf{w}_{r,m}^T \otimes \mathbf{\Theta}) \in \mathbb{C}^{Q\times MN}$, Eq. (6) is transformed into

$$\mathbf{y}_{r,m}^p = \mathbf{\Psi}_{r,m} \mathbf{h}_{r,m} + \mathbf{n}_{r,m}. \tag{7}$$

III. Proposed Framework

Let $\mathcal{D}_k$ represent the local dataset of a generic kth user, i.e. the input-output pair made of the received pilot signal ($\mathbf{y}_{r,m}^p$) and the channel vector ($\mathbf{h}_{r,m}$). The goal is to learn a parametric representation $f(\cdot|\boldsymbol{\theta})$ of the nonlinear input-output mapping with a decentralized methodology.

Local datasets are not transferred to the BS, but used at user location for training local models which are sent to the BS and aggregated into the global model ($\boldsymbol{\theta}$). The BS solves an optimization problem with a FL approach [19], i.e.

$$\min_{\boldsymbol{\theta}} \frac{1}{|\mathcal{U}|} \sum_{k\in\mathcal{U}} \mathcal{L}_k(\boldsymbol{\theta}; \mathcal{D}_k), \tag{8}$$

where $\mathcal{U}$ is the set of users and $\mathcal{L}_k(\cdot;\cdot)$ is the local objective function. In the tth communication round, the BS shares the global model ($\boldsymbol{\theta}_t$) with a subset of users ($\mathcal{U}_t \subset \mathcal{U}$). Based on the local objective function, the generic user updates the model using the local dataset via mini-batch or full-batch gradient descent methods (e.g., $\boldsymbol{\theta}_{k,t} \leftarrow \boldsymbol{\theta}_t - \eta \nabla \mathcal{L}_k(\boldsymbol{\theta}_t; \mathcal{D}_k)$ where η is the learning rate and $\boldsymbol{\theta}_{k,t}$ is the local model updated by the kth user at the tth communication round). Updated local models are aggregated by the BS as

$$\boldsymbol{\theta}_{t+1} = \sum_{k\in\mathcal{U}_t} \left(\frac{|\mathcal{D}_k|}{\sum_{j\in\mathcal{U}_t} |\mathcal{D}_j|}\right) \boldsymbol{\theta}_{k,t}, \tag{9}$$

and the procedure is repeated until convergence.

FL encounters significant challenges in the case of heterogeneous scenarios, which can significantly degrade performance [20]. The following steps are proposed for a communication-efficient robust framework for FL-based channel estimation.

A. Unsupervised Clustering

We propose a PCA-based clustering algorithm that identifies clusters solely based on the received pilot signals. Data for each used are assumed to be Z-score normalized and processed via standard PCA. Each user computes its own vector of variance ratios[2] (VRs) and share it with the BS.

[1]CSI of direct channels between the users and the BS is aquired by deactivating all RIS elements, akin to traditional massive MIMO systems [9].

[2]The entry of the VR vector is defined as the ratio between the single eigenvalue and the sum of all the eigenvalues.

The key assumption is that users belonging to the same region will experience similar channel conditions, which is reflected into similar VR vectors. K-means algorithm on VR vectors at the BS station, where the optimal number of clusters/regions is selected via the elbow method, allows for user clustering & region identification without prior knowledge.

B. FL & Adversary Identification

The proposed framework exploits clusters information to employ communication-efficient FL while discarding unsuitable models. In each communication round, the BS selects a subset of users from each region using cluster information and ensuring that contributions from all the regions are represented. Selected users receive the model from the BS and are requested to refine it according to their local dataset. We assume that local updates happen over ($\mathcal{E}$) epochs, with batch size B and using the MSE loss function ($\mathcal{L}_u(\cdot;\cdot)$) as local objective.[3] It is worth noticing that user selection can be based on different resource availability, thus allowing for effective handling of heterogeneous scenarios.

Model validation check (MVC) is conducted during the training process to prevent the incorporation of compromised model updates potentially resulting from adversarial users. In addition to the model parameters, each active user shares its local validation loss with the BS. A model is considered valid if the reported loss in the current round is close to the loss from the previous round, while deemed invalid and current update discarded if a significant increase in loss is observed.[4] The binary vector $\boldsymbol{\rho}$ collects the information about users' reliability based on the frequency of occurrences, with $\rho_k = 0$ (resp. $\rho_k = 1$) denoting the kth user is "trusted" (resp. "adversary").

User heterogeneity in terms of available resources is considered by letting users to train and share fewer parameters in case of limited capabilities, thus ensuring active participation of all users in the training process.

C. Communication Overhead

The communication overhead in CL is given by the amount of data (received pilots and channel vectors) transmitted by each user to the BS for centrally learning the model. The communication overhead in FL is given by the total number of parameters exchanged between each user and the BS during all the communication rounds [5], [13]. It is worth noticing that only the active users at each communication round contribute to the communication overhead.

IV. Simulation Results

We consider a single-cell RIS-assisted wireless system with a $M = 16$-antenna BS serving $K = 15$ single-antenna users via a $N = 64$-element UPA RIS. The cell has $R = 3$ distinct regions, each exhibiting unique channel characteristics and accommodating $U_r = 5$ users. The system operates at 28 GHz. The channel matrix $\mathbf{G}$ from the BS to the RIS is modeled incorporating $L_G = 3$ propagation paths between the RIS and the BS. The gains follow the complex Gaussian distribution $\alpha_{\ell_1}^{G} \sim \mathcal{CN}(0,1)$, while the phase angles are uniformly distributed within $(-\pi/2, \pi/2)$. The elements of $\mathbf{\Psi}_{r,m}$ are selected from $\left\{-\frac{1}{\sqrt{Q}}, \frac{1}{\sqrt{Q}}\right\}$, choosing one discrete phase shifts at the RIS and the BS, and $\mathbf{\Psi}_{r,m}$. The signal-to-noise ratio (SNR) is $1/\sigma_n^2$. The channel vector $\mathbf{f}_{r,m}$ from the RIS to the mth user in the rth region is generated using the Saleh-Valenzuela channel model. The number of paths between the RIS and each user is $L_{r,m} = 3$. The complex channel gains are modeled as complex Gaussian random variables $\alpha_{l_2}^{r,m} \sim \mathcal{CN}(0,1)$, and the phase angles are uniformly distributed across $(-\pi/2, \pi/2)$. The elevation angles are segmented according to 3 different scenarios: **Scenario 1** (non-overlapping regions) with elevation angles segmented as $(-\pi/2, -\pi/4)$, $(-\pi/6, \pi/12)$, and $(\pi/6, 5\pi/12)$; **Scenario 2** (adjacent regions) with elevation angles segmented as $(-\pi/2, -\pi/6)$, $(-\pi/6, \pi/6)$, and $(\pi/6, \pi/2)$; **Scenario 3** (overlapping regions)with elevation angles segmented as $(-\pi/2, -\pi/12)$, $(-\pi/6, \pi/4)$, and $(\pi/6, \pi/2)$. Scenario 2 is assumed unless explicitly specified differently.

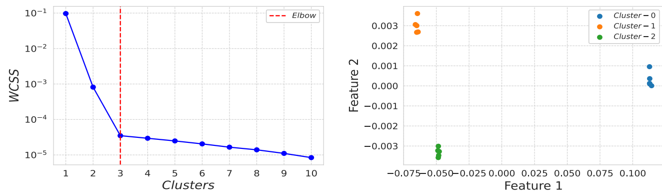
The pilot overhead is set to $Q = \frac{NM}{8} = 128$. Each user trains on $|\mathcal{D}_k| = 2 \times 10^4$ samples and employs the LS algorithm to generate labels for the cascaded channel (90% of its dataset for training and validation, 10% for testing) [9]. After Z-score normalization, PCA, and extraction of the VR vectors, K-means clustering is executed 10 times with different initializations and maximum number of iterations equal to 300.

Each user employs a layered architecture with approximately 2×10^6 trainable parameters including convolutional layers, batch-normalization (BN) layers, ReLU-activation layers, and fully-connected (FC) layers. The FL model is trained over 50 communication rounds and 25 communication rounds with multiple local epochs, using a learning rate of $\eta = 10^{-3}$, which is halved every 10-th iteration. ADAM optimizer with batch size $B = 256$ is used at user location. A small-model architecture with approximately 10^6 trainable parameters was considered to represent resource-constrained users.

The performance of the proposed algorithm is compared to classical FL-based channel estimation. **Strategy 1**: the model is trained using all users across all regions, running one single epoch per user for each communication round, namely All-User (AU) and explored in [5], [9], and [13]. **Strategy 2**: the model is trained randomly selecting 3 out of 15 users per communication round, running one single epoch. In this case, it is possible that one or two regions are not represented: to showcase the potential performance degradation when one region is missing. We trained a model that explicitly omits users from region 2, namely Missing-Region-2 (MR2). **Strategy 3**: the model is trained with the BS operating as Scenario 1 (resp. Scenario 2) for 60% (resp. 40%) of the communication rounds, namely MR2-40%. **Strategy 4**: the model is trained by smartly selecting users while safely integrating heterogeneous users, using single or multiple local epochs to save energy and reduce communication overhead, namely Efficient-User-Selection with m epochs (EUS-m).

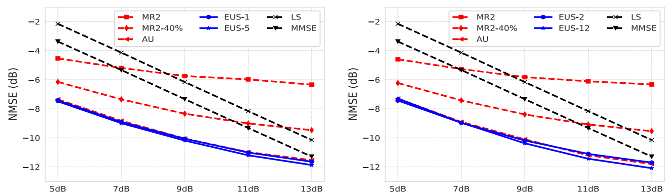
[3] The multiple-epoch refinement process facilitates faster convergence while reducing the number of communication rounds [19].

[4] A 2% increase in validation loss is allowed due to variability in user data and the stochastic nature of the training process.



(a) Elbow method (b) Clustering

Fig. 2. PCA-based clustering at 9 dB.



(a) K=15 (b) K=30

Fig. 3. NMSE performance comparison for SNR = {5, 7, 9, 11, 13} dB.

TABLE I
COMMUNICATION OVERHEAD

strategies	AU	CL	MR2	EUS-1	EUS-m
Floating Points ($\times 10^9$)	3.1	0.69	0.62	0.62	0.31

Fig. 2 shows the within-cluster sum of squares (WCSS) for different number of clusters, with 3 clusters being the optimal solution, and the corresponding user clustering in the subspace represented by the two main principal components.

Fig. 3 shows the normalized mean squared error (NMSE) of different strategies with SNR ranging from 5 dB to 13 dB on a test dataset representing the entire cell with equal contributions from each cluster. The proposed strategies (EUS-m) are compared with the traditional FL-based strategies (AU, MR2, and MR2-40%) and the additional benchmarks like the LS and MMSE algorithms for $K = 15$ and $K = 30$ users. It is apparent that the FL-based schemes outperform the traditional benchmarks. Although the performance gap narrows at higher SNR levels, where the LS and MMSE-based schemes achieve good estimation accuracy, they still demand significantly-higher pilot overhead.[5]

The results further highlight that selecting a smaller subset of users, especially in the multi-epoch scenario, substantially reduces the communication overhead while maintaining or slightly enhancing the performance. Table. I compares the communication overhead of different strategies. While the CL approach proves to be more communication efficient than the AU strategy, the proposed multi-epoch strategy significantly reduces the communication overhead and surpasses the CL approach. Moreover, EUS-1 strategy has a similar communication overhead as to the MR strategy (which fails to guarantee the inclusion of at least one user from each region).

Fig. 4 analyzes the performance of the proposed framework in a heterogeneous scenario where the users from one region are limited in their training capabilities and share only half of their model parameters. In the strategy denoted "small model", all users across all regions adopt a reduced model, thus enabling also resource-constrained users to participate in the training. In the EUS-1(Mixed Model (MM)) strategy, each user selects a full or reduced model according to its capabilities, again enabling all users to participate in the training. The strategies MR2 and EUS-1 are added as upper and lower bounds, respectively. It is apparent how the "small model" has similar degraded performance as MR2, while EUS-1(MM) is much closer to EUS-1, enabling effective participation of all users even in the case of heterogeneous scenarios.

Fig. 5 shows the algorithm's resilience under varying segmentation scenarios for the elevation angle. The results are consistently positive, irrespective of the region definition.

Fig. 6 shows the effectiveness of the proposed framework in presence of adversarial users. The proposed secured

[5]The pilot overhead for the LS and MMSE techniques is 1024, eight times higher than the pilot overhead of 128 needed for the FL-based schemes.

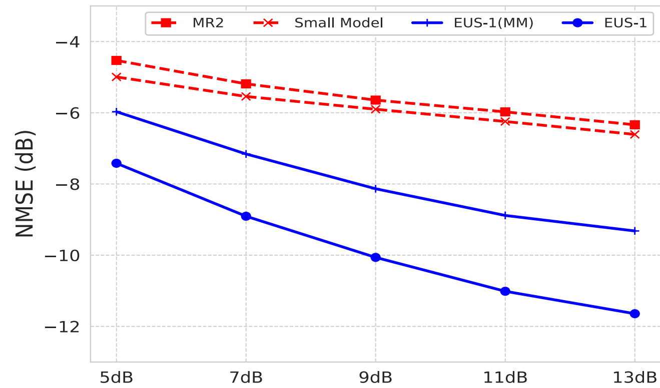


Fig. 4. Heterogeneous environment performance for SNR = {5, 7, 9, 11, 13} dB.

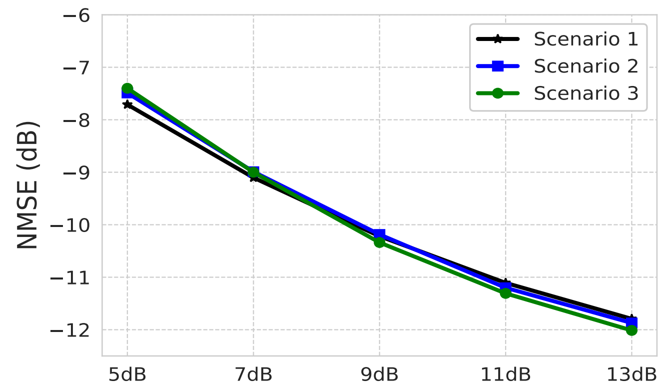


Fig. 5. Performance under different region scenarios.

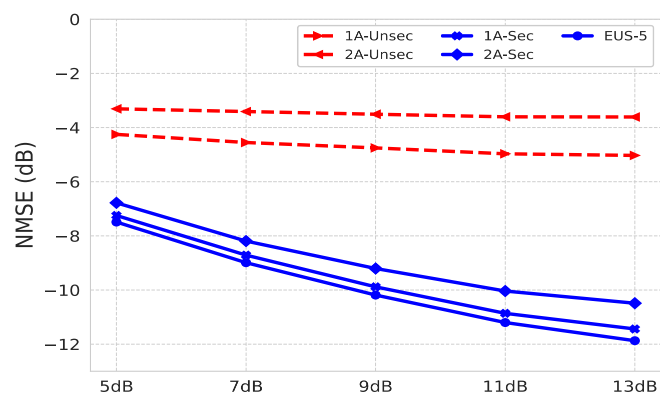


Fig. 6. Comparative analysis of secured vs. unsecured.

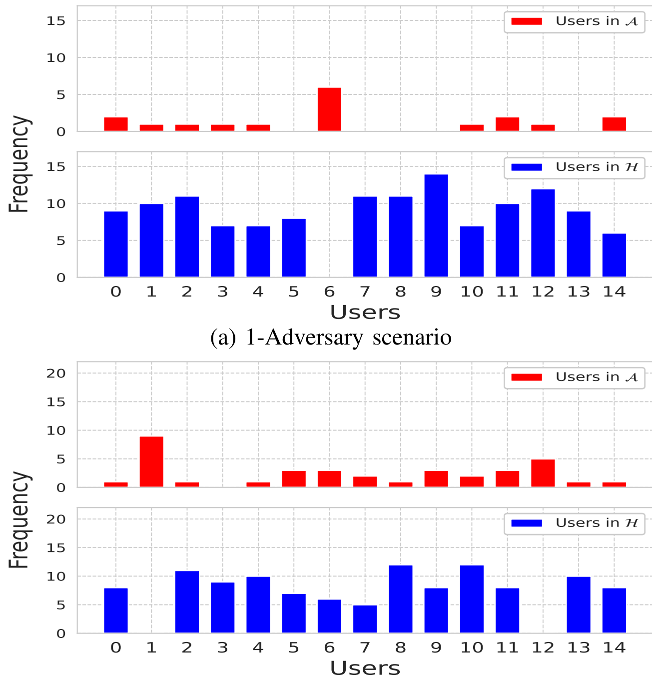


Fig. 7. Adversary identification.

approaches, namely 1A-Sec and 2A-Sec (involving 1 and 2 adversarial users) are compared with corresponding unsecured versions lacking MVC. The proposed secured approaches exhibit resilience in both adversarial scenarios and perform similar to the reference multi-epoch scenario without adversarial users. In contrast, the unsecured approaches experience significant performance degradation. To showcase the capability of user identification, Fig. 7 shows the frequency of users failing the MVC and flagged. User-6 acts as an adversary in the single-adversary scenario; User-1 and User-12 act as adversaries in the two-adversary scenario. It is apparent in Fig. 7a that User-6 exhibits the highest frequency of occurrences in the set of adversarial users ($\mathcal{A}$), highlighting its disruptive behavior. This contrasts with the users in the healthy set ($\mathcal{H}$), who passed the MVC, further confirming User-6's adversarial role. Similarly, in the two-adversary scenario, User-1 and User-12 show higher frequencies in the adversarial rounds, with no occurrences in rounds where the MVC is passed, emphasizing their roles as adversaries.

V. Conclusion

This letter presented a novel framework for FL-based channel estimation that reduces the communication overhead while enhancing the reliability and security of the FL framework in RIS-aided wireless systems. PCA-based unsupervised clustering is proposed to smartly selecting a subset of users from each cluster during each communication round. Numerical results showed that carefully selecting a limited number of users and adjusting the number of local training epochs can significantly reduce the communication overhead. The framework is specifically designed for effectively accommodating heterogeneous users by adjusting the model size according to their local resources. MVC was integrated to enhance system security against adversarial users corrupting the model parameters.

References

[1] J. Huang et al., "Reconfigurable intelligent surfaces: Channel characterization and modeling," *Proc. IEEE*, vol. 110, no. 9, pp. 1290–1311, Sep. 2022.

[2] M. D. Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[3] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1376–1393, Feb. 2021.

[4] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[5] B. Qiu, X. Chang, X. Li, H. Xiao, and Z. Zhang, "Federated learning-based channel estimation for RIS-aided communication systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 8, pp. 2130–2134, Aug. 2024.

[6] J. An, C. Xu, L. Gan, and L. Hanzo, "Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1245–1260, Feb. 2022.

[7] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4659–4663.

[8] Q. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, 2020.

[9] L. Dai and X. Wei, "Distributed machine learning based downlink channel estimation for RIS assisted wireless communications," *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4900–4909, Jul. 2022.

[10] J. An et al., "Codebook-based solutions for reconfigurable intelligent surfaces and their open challenges," *IEEE Wireless Commun.*, vol. 31, no. 2, pp. 134–141, Feb. 2024.

[11] M. Xu, S. Zhang, C. Zhong, J. Ma, and O. A. Dobre, "Ordinary differential equation-based CNN for channel extrapolation over RIS-assisted communication," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1921–1925, Jun. 2021.

[12] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9223–9228, Aug. 2020.

[13] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4255–4268, Jun. 2021.

[14] W. Shen, Z. Qin, and A. Nallanathan, "Federated learning enabled channel estimation for RIS-aided multi-user wireless systems," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 812–817.

[15] S. Abdulrahman, H. Tout, A. Mourad, and C. Talhi, "FedMCCS: Multicriteria client selection model for optimal IoT federated learning," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4723–4735, Mar. 2020.

[16] W. Chen, C.-K. Wen, X. Li, and S. Jin, "Adaptive bit partitioning for reconfigurable intelligent surface assisted FDD systems with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2488–2505, Apr. 2022.

[17] B. Guo, C. Sun, and M. Tao, "Two-way passive beamforming design for RIS-aided FDD communication systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.

[18] C. Hu, L. Dai, T. Mir, Z. Gao, and J. Fang, "Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8954–8958, Sep. 2018.

[19] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, vol. 54, 2017, pp. 1273–1282.

[20] A. Ghosh, J. Chung, D. Yin, and K. Ramchandran, "An efficient framework for clustered federated learning," in *Proc. Int. Conf. Neural Inf. Proc. Sys. (NeurIPS)*, 2020, pp. 19586–19597.