Physical Communication 71 (2025) 102698

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Full length article

Region-specific reliable channel estimation in RIS-enabled wireless communications via clustered federated learning ☆

Muhammad Asaad Cheema [a], *, Apoorva Chawla [a], Vinay Chakravarthi Gogineni [b], Pierluigi Salvo Rossi [a]

[a] Department of Electronic Systems, Norwegian University of Science and Technology, Høgskoleringen 1, Trondheim, 7034, Norway

[b] The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Campusvej 55, Odense M, DK-5230, Denmark

ARTICLE INFO

Keywords:
Clustered federated learning (CFL)
Autoencoder (AE)
Blockchain
Reconfigurable Intelligent Surfaces (RIS)

ABSTRACT

Machine learning (ML)-based downlink channel estimation for reconfigurable intelligent surface (RIS)-assisted communication faces challenges such as handling channel variations, high communication overhead of centralized learning (CL), and vulnerability to malicious users. We propose a novel approach integrating blockchain to enhance security by verifying registered users, autoencoder (AE)-based clustering to identify regions within the cell, and clustered federated learning (CFL) to ensure good channel estimation performance while minimizing communication and energy overhead. Simulations show that the proposed clustering-based scheme achieves estimation performance comparable to CL while significantly reducing communication and energy overhead.

1. Introduction

Reconfigurable intelligent surfaces (RIS) are among the key enablers of the advancements shaping the next-generation wireless systems [1–3]. RIS consists of multiple small, low-cost, passive reflective elements, each intelligently controlled to impose independent phase shifts on incident signals [4–7]. When strategically deployed, RIS mitigates line-of-sight challenges, particularly in terahertz (THz) and millimeter-wave (mmWave) communications [8,9]. However, realizing the full potential of RIS-enabled communication systems requires accurate downlink channel state information (CSI) [10], which poses a significant challenge due to the complexity of separately estimating channels from the base station (BS) to RIS and from RIS to the users. The complexity arises from the passive nature of RIS elements, requiring the estimation of the cascaded channel: while traditional techniques like least squares (LS) or minimum mean squared error (MMSE) can be used for cascaded channel estimation, the pilot overhead increases substantially in presence of RIS [11].

Machine learning (ML) techniques have effectively addressed channel estimation challenges: deep residual learning (DReL) was exploited to recover channel coefficients by learning residual noise [12]; CNN-based methods were considered to map low-dimensional cascaded channels to high-dimensional ones using a subset of RIS elements [13]; denoising NN-based cascaded channel estimation techniques were introduced in [14]. Although ML-based methods outperform traditional techniques, they rely on centralized frameworks requiring data transfer to a central server leading to high communication overhead and data integrity concerns [15]. Communication overhead can be reduced via federated learning (FL) where only model parameters (and not users data) are shared [16].

Despite the potential benefits, conventional FL-based approaches face challenges related to a single model not performing effectively across the whole cell due to variations in channel characteristics as users move between regions. Cell partitioning based on prior knowledge was explored in [11], but apparently limited by prior identification of the regions and does not include security mechanisms against malicious users. In our previous work, clustering was exploited for effective user selection, but the approach remained focused on building a single model [17].

This paper proposes a framework for downlink channel estimation in RIS-enabled communication systems based on clustered federated learning (CFL). More specifically, the proposed algorithm includes: (i) an autoencoder (AE)-based unsupervised region identification algorithm to group users based on the underlying characteristics of their received pilot signals, eliminating the need for prior knowledge; (ii) a simplified model-parameter selection which does not require an explicit classifier training to identify user transitions between regions; and (iii) a blockchain-based authentication mechanism that enhances security

☆ This work was supported in part by the Research Council of Norway through the Project ML4ITS within the IKTPLUSS Framework.

* Corresponding author.
E-mail address: asaad.cheema@ntnu.no (M.A. Cheema).

https://doi.org/10.1016/j.phycom.2025.102698
Received 20 March 2025; Received in revised form 15 April 2025; Accepted 21 April 2025
Available online 8 May 2025

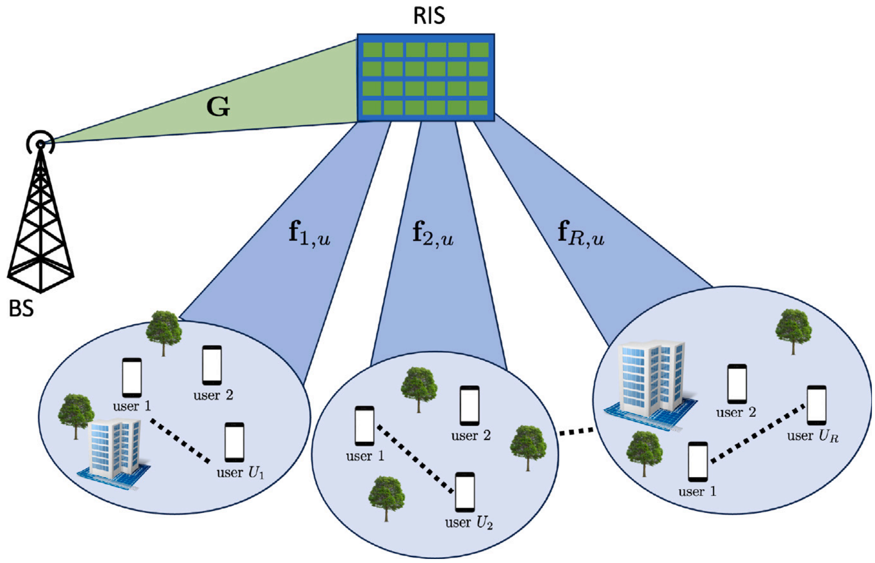


Fig. 1. Wireless communication through RIS assistance.

by ensuring only trustworthy users already connected to the BS participate in CFL training. Simulation results demonstrate the effectiveness of the unsupervised CFL in reducing communication and computational overhead without compromising system accuracy. Additionally, we investigate the impact of single-user poisoning within a cluster.

Notation. Uppercase (resp. lowercase) bold letters denote matrices (resp. column vectors). $(\cdot)^T$, $\otimes$, and $\|\cdot\|_2$ denote the transpose operator, the Kronecker product, and the Euclidean norm, respectively. $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix with $\mathbf{a}$ on the main diagonal. The gradient of a function is denoted by $\nabla f(\cdot)$. Calligraphic letters denote sets (except $\mathcal{L}$ denoting the loss function). $|\mathcal{C}|$ is the cardinality of $\mathcal{C}$.

2. System model

2.1. Channel model

We consider a single-cell, narrowband, RIS-aided, wireless communication system, where the BS communicates with K single-antenna users using an M-antenna uniform planar array (UPA). A RIS with N reflecting elements is positioned between the BS and the users to improve communications and the BS controls the RIS through a dedicated wireless link [9,18,19]. The cell is segmented into R regions, see Fig. 1, each accommodating a specific number of users (U_r such that $\sum_{r=1}^{R} U_r = K$). Each region exhibits different specific channel properties tailored to its users and represents a different channel scenario influenced by factors such as the distance between the users and the BS, the angles of arrival/departure of the channels, etc. The transition from one region to another may cause significant changes in the channel characteristics experienced by the user [11].

We consider the frequency division duplex (FDD) mode for the RIS-aided system, similar to [20,21], where the downlink and uplink channels exhibit non-reciprocity.[1] Note that we focus solely on obtaining the CSI for the reflecting link. Meanwhile, the direct channels between the BS and the users can be estimated by turning off all RIS elements,[2] similar to conventional massive MIMO systems [11]. Neglecting the effect of the direct channel and assuming that the BS communicates with the users via the RIS reflecting link, the downlink signal ($y_{r,u}$) received at the uth user in the rth region is

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

where $\mathbf{G} \in \mathbb{C}^{N\times M}$ represents the channel between the BS and the RIS, $\mathbf{f}_{r,u} \in \mathbb{C}^{N\times 1}$ denotes the channel between the RIS and the uth user in the rth region, and $n_{r,u} \in \mathbb{C}$ represents the additive white Gaussian noise. Further, $\mathbf{w}_{r,u} \in \mathbb{C}^{M\times 1}$ and $x_{r,u} \in \mathbb{C}$ denote the precoding vector at the BS and the transmitted signal by the BS, respectively. The reflecting matrix at the RIS, denoted by $\mathbf{\Phi} \in \mathbb{C}^{N\times N}$, is diagonal, with $\mathbf{\Phi} = \mathrm{diag}(\phi_1, \phi_2, \dots, \phi_N)$, where ϕ_n is the reflection coefficient of the nth element of the RIS. Using the Saleh-Valenzuela model, the channel $\mathbf{G}$ can be expressed as

$$\mathbf{G} = \sqrt{\frac{MN}{L_G}} \sum_{l_1=1}^{L_G} \alpha_{l_1}^{G} \mathbf{a}\left(\vartheta_{l_1}^{G_\mathrm{r}}, \psi_{l_1}^{G_\mathrm{r}}\right) \mathbf{b}\left(\vartheta_{l_1}^{G_\mathrm{t}}, \psi_{l_1}^{G_\mathrm{t}}\right)^T , \tag{2}$$

where $\alpha_{l_1}^{G}$ signifies the complex gain, $\vartheta_{l_1}^{G_\mathrm{r}}(\psi_{l_1}^{G_\mathrm{r}})$ and $\vartheta_{l_1}^{G_\mathrm{t}}(\psi_{l_1}^{G_\mathrm{t}})$ denote the azimuth (elevation) angles at the RIS and the BS for the l_1th path, respectively, and L_G is the number of paths between the RIS and the BS. Further, $\mathbf{G}$ remains nearly constant once the RIS is installed. Likewise, the channel $\mathbf{f}_{r,u} \in \mathbb{C}^{N\times 1}$ can be given as

$$\mathbf{f}_{r,u} = \sqrt{\frac{N}{L_{r,u}}} \sum_{l_2=1}^{L_{r,u}} \alpha_{l_2}^{r,u} \mathbf{a}\left(\vartheta_{l_2}^{r,u}, \psi_{l_2}^{r,u}\right) , \tag{3}$$

where $L_{r,u}$ is the number of paths between the uth user in the rth region and the RIS. The variables $\alpha_{l_2}^{r,u}$, $\vartheta_{l_2}^{r,u}$, and $\psi_{l_2}^{r,u}$ denote the complex gain, the azimuth angle, and the elevation angle associated with the l_2 path of the RIS, respectively. Further, $\mathbf{b}(\vartheta, \psi) \in \mathbb{C}^{M\times 1}$ and $\mathbf{a}(\vartheta, \psi) \in \mathbb{C}^{N\times 1}$ correspond to the normalized steering vectors linked with the BS and the RIS, respectively. For a standard $N_1 \times N_2$ UPA, the expression for $\mathbf{a}(\vartheta, \psi)$ can be given as [26]

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

where $\eta = 2\pi d/\lambda$, $N = N_1 N_2$, $\mathbf{n}_1 = [0, 1, \dots, N_1 - 1]^T$ and $\mathbf{n}_2 = [0, 1, \dots, N_2 - 1]^T$, λ denotes the carrier wavelength and the antenna spacing d is typically chosen as $d = \lambda/2$. Defining $\boldsymbol{\phi} = [\phi_1, \phi_2, \dots, \phi_N]^T \in \mathbb{C}^{N\times 1}$, the system model in (1) can be remodeled as

$$\begin{aligned} y_{r,u} &= \boldsymbol{\phi}^T \mathrm{diag}(\mathbf{f}_{r,u}^T) \mathbf{G} \mathbf{w}_{r,u} x_{r,u} + n_{r,u}, \\ &= \boldsymbol{\phi}^T \mathbf{H}_{r,u} \mathbf{w}_{r,u} x_{r,u} + n_{r,u}, \end{aligned} \tag{5}$$

[1] Even in the time division duplex (TDD) mode, the uplink and downlink channels may not be reciprocal in certain hardware implementations of RIS, as explicitly stated in [22]. Accurate downlink channel estimation is required to achieve effective beamforming for both the BS and the RIS [23].

[2] A practical method for turning off the RIS is proposed in [24,25].

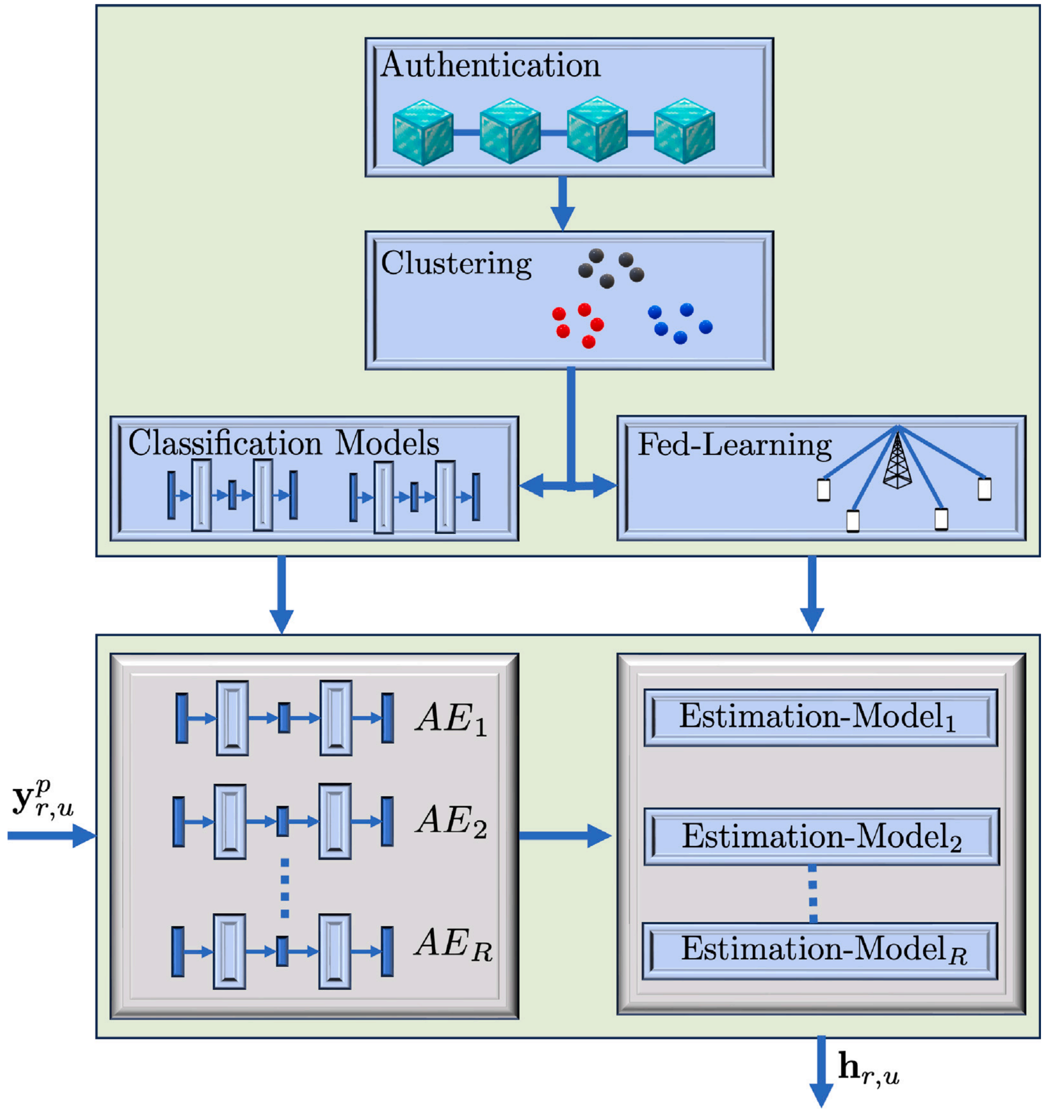


Fig. 2. Proposed framework for cascaded downlink channel estimation.

where the downlink cascaded channel $\mathbf{H}_{r,u} \in \mathbb{C}^{N\times M}$ for the uth user in the rth region is defined as $\mathbf{H}_{r,u} = \text{diag}(\mathbf{f}^T_{r,u})\mathbf{G}$. Due to the inability of the RIS to process the signals independently, the cascaded channel $\mathbf{H}_{r,u}$ is typically estimated rather than distinct channels $\mathbf{G}$ and $\mathbf{f}_{r,u}$.

For the downlink cascaded channel estimation, the BS transmits known pilot signals to the users through the RIS over Q time slots. Using (5), the pilot signal received at the uth user in the rth region during the qth time slot, denoted by $y^p_{r,u,q} \in \mathbb{C}$, can be expressed as

$$y^p_{r,u,q} = \boldsymbol{\phi}^T_q \mathbf{H}_{r,u}\mathbf{w}_{r,u} p_{r,u,q} + n_{r,u,q}, \tag{6}$$

where $q = 1, 2, \dots, Q$. Further, $p_{r,u,q}$, $\boldsymbol{\phi}_q$ and $n_{r,u,q} \sim \mathcal{CN}(0, \sigma^2_n)$ denote the pilot signal transmitted by the BS, the reflection vector at the RIS and the noise at the user associated with the qth time slot, respectively. Considering Q time slots, the received pilot vector $\mathbf{y}^p_{r,u}$ at the uth user can be given as

$$\mathbf{y}^p_{r,u} = \boldsymbol{\Theta}\mathbf{H}_{r,u}\mathbf{w}_{r,u} + \mathbf{n}_{r,u}, \tag{7}$$

where $\mathbf{y}^p_{r,u} = [y^p_{r,u,1}, y^p_{r,u,2}, \dots, y^p_{r,u,Q}]^T \in \mathbb{C}^{Q\times 1}$, the pilot signal is chosen as $p_{r,u,q} = 1$, $\boldsymbol{\Theta}^T = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_Q] \in \mathbb{C}^{N\times Q}$ and $\mathbf{n}_{r,u} = [n_{r,u,1}, n_{r,u,2}, \dots, n_{r,u,Q}]^T \in \mathbb{C}^{Q\times 1}$. Using the identity $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^\mathbf{T} \otimes \mathbf{A})\text{vec}(\mathbf{B})$, the above system model can be modified as

$$\begin{aligned}\mathbf{y}^p_{r,u} &= (\mathbf{w}_{r,u}{}^T \otimes \boldsymbol{\Theta})\text{vec}(\mathbf{H}_{r,u}) + \mathbf{n}_{r,u},\\ &= \boldsymbol{\Psi}_{r,u}\mathbf{h}_{r,u} + \mathbf{n}_{r,u},\end{aligned} \tag{8}$$

where $\boldsymbol{\Psi}_{r,u} = (\mathbf{w}_{r,u}{}^T \otimes \boldsymbol{\Theta}) \in \mathbb{C}^{Q\times MN}$ and $\mathbf{h}_{r,u} = \text{vec}(\mathbf{H}_{r,u}) \in \mathbb{C}^{MN\times 1}$. It is assumed that $\boldsymbol{\Theta}$ and $\mathbf{w}_{r,u}$ remain fixed and are known to both the BS and the user. The primary goal is to estimate $\mathbf{h}_{r,u}$ using $\mathbf{y}^p_{r,u}$ and $\boldsymbol{\Psi}_{r,u}$.

3. Proposed algorithm

The proposed framework consists of two main blocks (see Fig. 2). The upper block includes authentication, training, classification, and CFL-based channel estimation models. The lower block leverages these models to categorize input pilot signals into specific regions, and subsequently employs the designated region models for channel estimation, akin to the testing phase.

3.1. Blockchain-based authentication

Authentication is implemented to establish a secure environment and ensure the participation of trustworthy users only. In traditional

FL-based schemes for CSI estimation, the users are typically authenticated to confirm their registration and connectivity with the BS. However, this process could allow adversaries to connect to the BS and poison the model parameters merely by participating in the network. To address this vulnerability, a blockchain-based authentication mechanism has been implemented on top of the traditional system for CFL training, thus trustworthy users participate in the training, while other users receive the final trained model for channel estimation purposes. Blockchain technology is selected due to four crucial benefits: (i) data integrity through cryptographic features, (ii) user privacy, (iii) distributed structure robust to single point of failure, and (iv) public records for fair and transparent auditing [27–29].

A smart contract on the Ethereum platform handles user authentication and registration performing two main functions: (i) registering trustworthy users based on the hash of their identity, and (ii) authenticating users before they participate in CFL model training. Each trustworthy user is assigned a unique identifier for registration, such as an international mobile subscriber identifier (IMSI), coupled with a secret phrase, both managed by a designated authority.[3] The authority is the only entity authorized to update the list of registered users. When a user connects to the BS, the BS hashes the IMSI along with the secret phrase to confirm the user's registration. Since this is a one-time process, sharing the secret imposes minimal to no overhead. Once registration is confirmed, the user can participate in the CFL-based training. Furthermore, the blockchain is managed by the BS, which has sufficient computational resources to handle them efficiently.

3.2. Unsupervised clustering and classification

Each authenticated user trains a lightweight AE using its received signals for a limited number of epochs. It is crucial to note that at this stage, the BS and the users operate without prior knowledge of the regions. Each user subsequently shares its encoder/decoder model parameters and 5% of its received pilot signals with the BS. The BS tests data from each user against all AE parameters and generates a loss matrix where each row serves as user embedding, capturing the loss values corresponding to user's data when processed by all other users' AEs. The embeddings are then clustered using K-means algorithm, where the optimal number of clusters is determined using the elbow method.[4] This regional information is then utilized to train tailored region-specific models.

The BS selects one user per cluster[5] and assumes the AE parameters of the selected user as representative for the corresponding region. The set of selected AE parameters (one per region) allows for effective assignment of channel-estimation models to mobile users, who can potentially transit from one region into a different region, and to new users joining the system. The effective assignment relies on classification based on the minimum reconstruction loss of the reference AE parameters operating on signals from the user of interest. It is worth noticing that the classifier is built on the classes identified by the clustering algorithm and thus requires no prior information. The overall methodology of this clustering process is detailed in Algorithm 1.

Algorithm 1 AE-based Clustering and Classification

procedure TRAIN-AE
  for kth user, $k = 1, \dots, K$, **do**
    Initialize $\boldsymbol{\beta}, \boldsymbol{\gamma}$
    $\min_{\gamma,\beta} \left(\frac{1}{s} \sum_{i=1}^{s} \|\mathbf{x}_{k,i} - \hat{\mathbf{x}}_{k,i}\|_2^2 \right)$
    $AE_k = [\boldsymbol{\beta}, \boldsymbol{\gamma}]$
    Share AE_k and $\mathcal{D}_k^{\mathrm{AE}}$ with the BS
  end for
end procedure
procedure LOSS MATRIX (LM)
  Initialize LM with zeros
  for each AE_i, $i = 1, \dots, K$, **do**
    for each $\mathcal{D}_j^{\mathrm{AE}}$, $j = 1, \dots, K$, **do**
      $LM_{i,j} = \mathcal{L}(\mathbf{X}, \hat{\mathbf{X}}), \text{ where } \begin{cases} \mathbf{X} \in \mathcal{D}_j^{\mathrm{AE}}, \\ \hat{\mathbf{X}} = AE_i(\mathbf{X}) \end{cases}$
    end for
  end for
end procedure
procedure CLUSTERING
  $\mathcal{Z} = \mathrm{EM}(LM)$ ▷ EM represents the Elbow Method.
  $Clts = \mathrm{KM}(LM, \mathcal{Z})$ ▷ $Clts = \{C_1, C_2, \dots, C_{\mathcal{Z}}\}$ is the set of clusters, and KM represents the K-means algorithm.
end procedure
procedure CLASSIFICATION
  $AE_{\text{Classify}} = \{f(C_1), f(C_2), ..., f(C_{\mathcal{Z}})\}$ ▷ AE_{Classify} represents the AE-based classification of the received signals, and $f(C_i)$ selects an AE from cluster C_i.
  $Loss = [\ ]$
  for each AE in AE_{Classify} **do**
    $\mathcal{L}_{AE} = \mathcal{L}(\mathbf{y}^p, AE(\mathbf{y}^p))$
    $Loss \leftarrow Loss \cup \{\mathcal{L}_{AE}\}$
  end for
  $I = Index(\min(Loss))$
end procedure

3.3. Communication and energy efficient training

After identifying different regions within a cell, the CFL optimizes the parameters for each identified region/cluster using the collective contributions of the corresponding users. During each communication round, the BS selects a subset ($\mathcal{S}$) of users from a region and provides region-specific model parameters to solve the following optimization problem [33]:

$$\min_{\theta_r} \underbrace{\frac{1}{|\mathcal{S}|} \sum_{u \in \mathcal{S}} \mathcal{L}_k(\theta_r; \mathcal{D}_u)}_{\mathcal{L}(\theta; \mathcal{D})}, \tag{9}$$

where the vector θ_r collects the globally shared model parameters within a region and $\mathcal{L}_k(\theta; \mathcal{D}_u)$ is the local objective function used in conjunction with the local dataset $\mathcal{D}_u$ for the uth client. The users refine the parameters using their local datasets over multiple epochs with a batch size of B and send the updated parameters back to the BS. The BS then aggregates the refined models and updates the associated cluster models. This iterative process continues until convergence or a predetermined number of communication rounds is reached.

Unlike traditional approaches [11,16], the proposed framework involves only a subset of users per communication round, significantly reducing communication load and the BS's computational burden for parameter aggregation. Random user selection further conserves energy by preventing continuous user activity. The CFL training algorithm is summarized in Algorithm 2.

[3] It refers to the administrative entity within the mobile subscriber company or the organization responsible for the framework's design. Furthermore, user trust for registration can be determined by analyzing network traffic and user behavior [30–32].

[4] Ideally, the number of clusters equals the number of regions (R).

[5] The impact of the user selection strategy (*e.g.* random selection, optimized strategy) falls beyond the scope of this work.

Algorithm 2 Clustered Federated Learning

procedure Federated Learning
  Initialize: Global model weights $\theta_{r,0}$
  for each round $n = 1, 2, \ldots$ **do**
    $S_n \leftarrow \text{RandomSelect}(U_r, m)$ ▷ Selecting m users from U_r users corresponding to the rth region
    for each uth user in S_n **do**
      $\theta_{r,u,n} \leftarrow \text{ClientUpdate}(\mathcal{D}_{r,u}, \theta_{r,n})$
    end for
    $\theta_{r,n+1} \leftarrow \frac{1}{|S_n|} \sum_{u \in S_n} \theta_{r,u,n}$
  end for
end procedure
procedure ClientUpdate$(\mathcal{D}_{r,u}, \theta_r)$
  $\mathcal{B}_k \leftarrow Batches(\mathcal{D}_{r,u}, B)$
  $\theta_{r,u} \leftarrow \theta_r$
  for $e = 1$ to $\mathcal{E}$ **do** ▷ $\mathcal{E}$ represents the number of epochs
    for each $b \in \mathcal{B}_k$ **do**
      $\theta_{r,u} \leftarrow \theta_{r,u} - \eta \nabla \mathcal{L}_{r,u}(\theta_{r,u}; b)$
    end for
  end for
  Share $\theta_{r,u}$ with the BS
end procedure

Table 1
Autoencoder layer architecture.

Layer type	Input size	Output size
Linear	256	512
Linear	512	256
Linear	256	128
Linear	128	64
Linear	64	32
Linear	32	16
Linear	16	32
Linear	32	64
Linear	64	128
Linear	128	256
Linear	256	512
Linear	512	256

Table 2
Personalized cluster model architecture.

Layer type	Filter size	Output channels
Convolutional Layer 1	3 × 3	32
Batch Normalization Layer	–	–
Convolutional Layer 2	3 × 3	16
Batch Normalization Layer	–	–
Convolutional Layer 3	3 × 3	8
Batch Normalization Layer	–	–
Fully Connected Layer	–	Output

3.4. Communication overhead

Communication overhead (O^C) refers to data exchange during the training phase of the model. The communication overhead in centralized learning (CL) is characterized by the number of data symbols transmitted by the user to the BS for training the centralized model, comprising both input samples and their corresponding outputs. Conversely, the communication overhead for CFL, is defined as the total number of parameters transmitted by the user and received from the BS during the CFL training [16].

3.5. Energy overhead

Energy overhead (O^E) refers to the number of iterations required to update the model parameters during training. The energy overhead in CL is the product of the number of batches and the number of epochs. Conversely, the energy overhead in FL is characterized by the product of the number of batches per user, the number of epochs per user, the number of selected users per communication round, and the total number of communication rounds.[6]

4. Simulation results

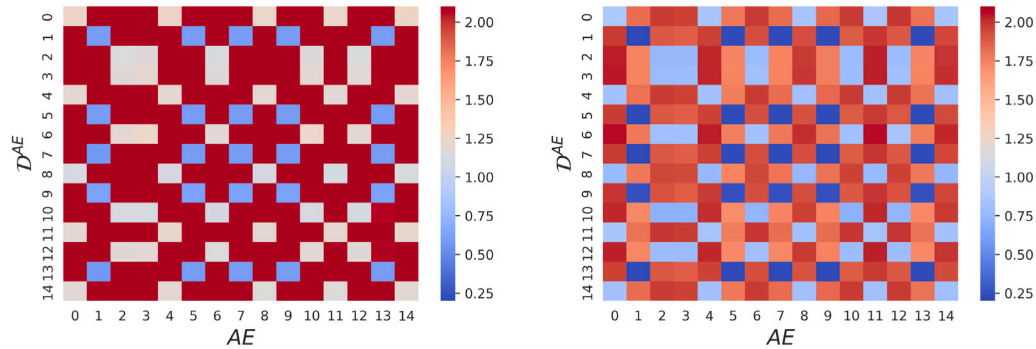
This section presents the simulation results to validate the performance of our proposed algorithms. We consider a single-cell, RIS-assisted communication system, where a BS with $M = 16$ antenna serves $K = 15$ single-antenna users using a $N = 64$ element UPA RIS. The cell is segmented into $R = 3$ regions, where each region has different channel characteristics and contains $U_r = 5$ users for $r = 1, 2, 3$. The channel matrix $\mathbf{G}$ between the BS and the RIS is generated using Eq. (2) with $L_G = 3$ paths between the BS and the RIS. The complex gains are modeled as complex Gaussian random variables with zero mean and unit variance, *i.e.* $\alpha^G_{l_1} \sim \mathcal{CN}(0, 1)$, and the phase angles are uniformly distributed over $(-\pi/2, \pi/2)$. Each element of $\boldsymbol{\Psi}_{r,u}$ in Eq. (8) is chosen from the set $\left\{-\frac{1}{\sqrt{Q}}, \frac{1}{\sqrt{Q}}\right\}$, considering discrete phase shifts at the BS and the RIS, and $\boldsymbol{\Psi}_{r,u}$ is assumed to be identical $\forall u, r$. The signal-to-noise ratio (SNR) is defined as $1/\sigma_n^2$, where σ_n^2 is the noise variance. We employ the Saleh-Valenzuela channel model for the channel $\mathbf{f}_{r,u}$ in Eq. (3) from RIS to the uth user in the rth region. The path count between the RIS and the user is $L_{r,u} = 3$, $\forall u, r$. The complex channel gains $\alpha^{r,u}_{l_2}$ follow the complex Gaussian distribution, *i.e.*, $\alpha^{r,u}_{l_2} \sim \mathcal{CN}(0, 1)$ and phase angles are uniformly distributed over $(-\pi/2, \pi/2)$. The elevation angles are equally divided into three segments $(-\pi/2, -\pi/6)$, $(-\pi/6, \pi/6)$, and $(\pi/6, \pi/2)$, with each segment corresponding to a specific region and channel characteristics. The operating frequency is 28 GHz.

The required pilot overhead is set as $Q = \frac{NM}{8} = 128$ [11]. Each user collects a dataset of $|\mathcal{D}_{r,u}| = 20{,}000$ samples, with 10% reserved for testing. Each user is equipped with an AE for clustering and classification. The received pilot signal is transformed into a real-valued vector by concatenating its real and imaginary parts, with the first half containing the real part and the second half containing the imaginary part. The AE's takes this as input and maps it to 512 features, then progressively reduces the feature size by half in each step until reaching 16. It then reconstructs the data by symmetrically increasing the feature size back to the original input dimensions. Table 1 outlines the AE architecture. The AE is trained using the Adam optimizer, with a learning rate of $\eta = 1 \times 10^{-3}$ and a batch size of $B = 1024$ in PyTorch.

Each personalized cluster model employs a CNN architecture with three 3 × 3 convolutional layers, producing 32, 16, and 8 channels, followed by batch normalization and ReLU activation. The output is then flattened and passed through a fully connected layer, transforming into a 2048-dimensional feature vector. The complete model includes approximately 2.1 million trainable parameters, as detailed in Table 2. We analyzed five scenarios to evaluate the impact of selecting the optimal number of users and local training epochs per communication round in CFL: in scenario S1, all five users in a cluster participate, each running one local training epoch; in scenario S2, three out of five users per round are selected, each training for one epoch; in scenario S3 two users per round are randomly selected, each training for one epoch; in scenario S4, two randomly chosen users train for three epochs each; and in scenario S5, a single user is selected to train for five epochs. The ML model architecture remains unchanged across all configurations.

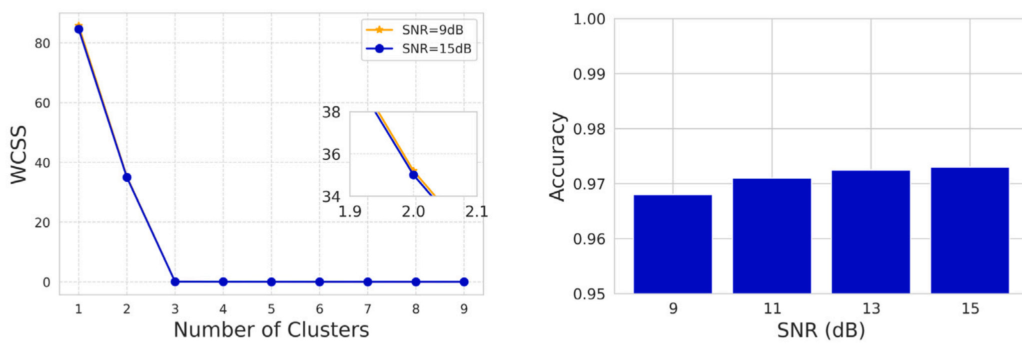
Fig. 3 presents the loss metrics by evaluating each user's data against AE models trained on individual users. The vertical axis represents tested users, while the horizontal axis corresponds to AE models.

[6] If all users participate in each round, FL experiences the same energy overhead as CL.



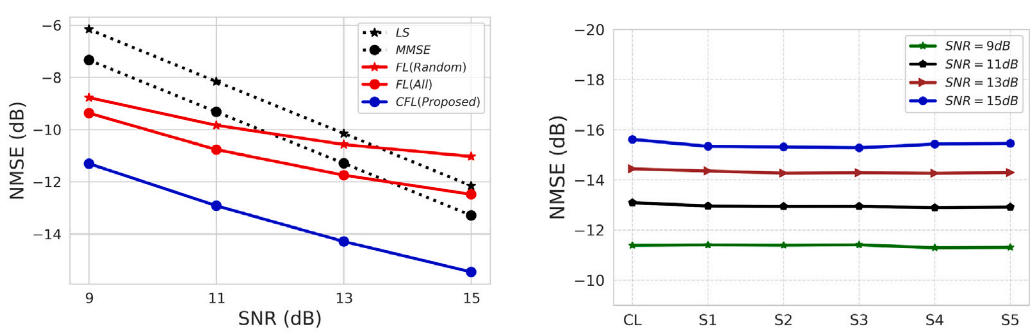
(a) SNR=9 dB (b) SNR=15 dB

Fig. 3. Loss grid.



(a) Optimal number of clusters (b) Classifier accuracy

Fig. 4. Clustering and classification performance.



(a) CFL vs. single-model FL (b) NMSE for different scenarios

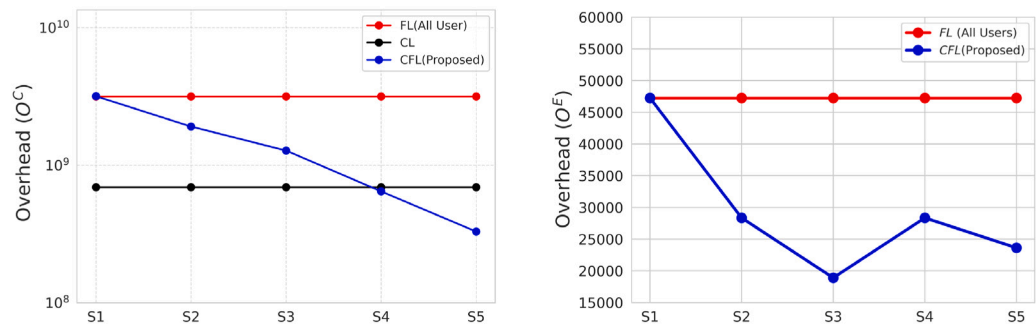
Fig. 5. Performance comparison.

The AE's ability to regenerate similar signal structures is utilized for clustering, where similar loss embeddings reflect shared user characteristics. As shown in Fig. 3, an AE model trained on one user's data typically yields lower reconstruction loss on data that shares similar statistical properties. In contrast, when the AE is applied to data from users in different clusters (*e.g.*, with differing environmental conditions), the reconstruction loss tends to be higher. Finally, the K-means algorithm, applied to the loss metrics, identifies user clusters, enabling the training of a cluster-specific model for each cluster.

The within-cluster sum of squares (WCSS) is calculated to determine the optimal number of clusters, as shown in Fig. 4(a). Lower WCSS values indicate tighter clusters with data points closer to their centroids. For the classifier, we consider random selection in each cluster. Fig. 4(b) shows the classifier's efficiency on 30,000 test samples, maintaining 97 ± 1% accuracy without additional training. This showcases the classifier's ability to correctly identify cluster and select the appropriate cluster-specific model for channel estimation.

Fig. 5(a) compares CFL with two traditional FL approaches both training a single global model: one based on randomly selecting three users from each region, and the other using all available users. Conversely, CFL trains region-specific models based on user clustering. Traditional LS and MMSE methods are also shown as benchmarks; the required pilot overhead for LS and MMSE is set at $Q = NM = 1024$. Apparently, the proposed approach significantly outperforms the traditional benchmarks and the conventional FL alternatives, demonstrating its superior performance.

Fig. 5(b) analyzes the performance of various strategies on a test dataset of 30,000 samples that cover the entire cell, ensuring equal contributions from each cluster. Testing involves classifying input samples into their regions before applying specific channel estimation models.



(a) Communication overhead. (b) Energy overhead.

Fig. 6. Comparison of communication and energy overhead.

Table 3
Secured vs. unsecured at SNR = 15 dB.

	S1	S2	S3	S4	S5
Secured (dB)	−15.55	−15.31	−15.28	−15.43	−15.40
Unsecured (dB)	−2.69	−1.43	−10.57	−12.86	−12.47

Results demonstrate that selecting a smaller subset of users within the cluster and fewer communication rounds maintains performance instead of utilizing all users in traditional FL and centralized methods for channel estimation. The proposed CFl based strategy even slightly enhances the performance, particularly in scenario S4 at 15 dB, compared to the all-users scenario.

Fig. 6(a) presents a comparative analysis of various training configurations, focusing on communication overhead of CL, conventional FL, and proposed CFL approaches. It is apparent that CL is more efficient than conventional FL in all scenarios and surpasses the proposed CFL in scenarios S1, S2, and S3, while the CFL[7] is preferable in scenarios S4 and S5. Fig. 6(b) the energy overhead during training for channel estimation, showing that effective user selection significantly impacts the overall energy consumption. Additionally, energy utilization per user decreases as the probability of selecting a specific user drops from 1 to 0.2 when considering a single user. Overall, Fig. 6 validate that the proposed CFL-based strategy achieves strong performance while significantly reducing communication and computational overhead, highlighting its effectiveness and efficiency for scalable deployment.

Table 3 compares two CFL systems: one with authentication, ensuring all five users are non-malicious, and another allowing all users, including one malicious user. The malicious user, once selected, receives the model from the BS, ignores the learning process, randomly reinitializes the parameters, and returns them to the BS. These actions disrupt training and worsen NMSE, especially in scenarios S1 and S2, where the adversarial user is frequently selected. Differently, in scenarios S4 and S5 the chances of excluding the adversarial user increase and system performance improves. Generally, the system performs optimally when all users are authenticated and trustworthy, ensuring accurate channel estimation.

5. Conclusion and future work

This paper proposed a CFL-based framework enhancing communication efficiency, energy efficiency, and security. We designed an unsupervised clustering and classification method based on the received pilot signals, which enhances the channel estimation accuracy since users within the same cluster can collaborate more efficiently. We demonstrated that strategically selecting fewer users and optimizing the number of local epochs can significantly improve communication efficiency and energy efficiency. Additionally, we employed blockchain technology to add a layer of protection and prevent malicious users from compromising the model. For future work, we aim to explore scenarios involving user mobility and various adversarial situations.

[7] Note that our approach accounts for the overhead of sharing 5% of data and AE parameters for clustering and classification.

CRediT authorship contribution statement

Muhammad Asaad Cheema: Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Apoorva Chawla:** Writing – review & editing, Writing – original draft, Conceptualization. **Vinay Chakravarthi Gogineni:** Writing – review & editing, Conceptualization. **Pierluigi Salvo Rossi:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Muhammad Asaad Cheema reports financial support was provided by Research Council of Norway. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] M. Di Renzo, F.H. Danufane, S. Tretyakov, Communication models for reconfigurable intelligent surfaces: From surface electromagnetics to wireless networks optimization, Proc. IEEE 110 (9) (2022) 1164–1209.

[2] G.C. Nwalozie, A.L. de Almeida, M. Haardt, Enhanced channel estimation for double RIS-aided MIMO systems using coupled tensor decompositions, Signal Process. (2025) 109979, http://dx.doi.org/10.1016/j.sigpro.2025.109979, URL https://www.sciencedirect.com/science/article/pii/S0165168425000933.

[3] Y. Sun, Z. Lin, K. An, D. Li, C. Li, Y. Zhu, D. Wing Kwan Ng, N. Al-Dhahir, J. Wang, Multi-functional RIS-assisted semantic anti-jamming communication and computing in integrated aerial-ground networks, IEEE Journal on Selected Areas in Communications 42 (12) (2024) 3597–3617.

[4] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, H.V. Poor, Active RIS vs. Passive RIS: Which will prevail in 6G? IEEE Trans. Commun. 71 (3) (2023) 1707–1725, http://dx.doi.org/10.1109/TCOMM.2022.3231893.

[5] C. Zhang, Y. Liu, B.K. Ng, C.-T. Lam, RIS-assisted differential transmitted spatial modulation design, Signal Process. 230 (2025) 109767, http://dx.doi.org/10.1016/j.sigpro.2024.109767, URL https://www.sciencedirect.com/science/article/pii/S0165168424003876.

[6] L. Zhi, N. Hehao, H. Yuanzhi, A. Kang, Z. Xudong, C. Zheng, X. Pei, Self-powered absorptive reconfigurable intelligent surfaces for securing satellite-terrestrial integrated networks, China Communications 21 (9) (2024) 276–291, http://dx.doi.org/10.23919/JCC.fa.2023-0437.202409.

[7] Z. Lin, H. Niu, K. An, Y. Wang, G. Zheng, S. Chatzinotas, Y. Hu, Refracting RIS-Aided Hybrid Satellite-Terrestrial Relay Networks: Joint Beamforming Design and Optimization, IEEE Transactions on Aerospace and Electronic Systems 58 (4) (2022) 3717–3724, http://dx.doi.org/10.1109/TAES.2022.3155711.

[8] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, S. Tretyakov, Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead, IEEE J. Sel. Areas Commun. 38 (11) (2020) 2450–2525.

[9] B. Ning, Z. Chen, W. Chen, Y. Du, J. Fang, Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming, IEEE Trans. Veh. Technol. 70 (2) (2021) 1376–1393.

[10] S. Li, R. Yang, X. Guo, G. Han, J. Zhang, Parametric channel estimation for RIS-assisted mmwave MIMO-OFDM systems with low pilot overhead, Signal Process. 225 (2024) 109611, http://dx.doi.org/10.1016/j.sigpro.2024.109611, URL https://www.sciencedirect.com/science/article/pii/S0165168424002305.

[11] L. Dai, X. Wei, Distributed machine learning based downlink channel estimation for RIS assisted wireless communications, IEEE Trans. Commun. 70 (7) (2022) 4900–4909.

[12] C. Liu, X. Liu, D.W.K. Ng, J. Yuan, Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications, IEEE Trans. Wirel. Commun. 21 (2) (2021) 898–912.

[13] M. Xu, S. Zhang, C. Zhong, J. Ma, O.A. Dobre, Ordinary differential equation-based CNN for channel extrapolation over RIS-assisted communication, IEEE Commun. Lett. 25 (6) (2021) 1921–1925.

[14] S. Liu, Z. Gao, J. Zhang, M. Di Renzo, M.-S. Alouini, Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces, IEEE Trans. Veh. Technol. 69 (8) (2020) 9223–9228.

[15] B. Qiu, X. Chang, X. Li, H. Xiao, Z. Zhang, Federated learning-based channel estimation for RIS-aided communication systems, IEEE Wirel. Commun. Lett. (2024).

[16] A.M. Elbir, S. Coleri, Federated learning for channel estimation in conventional and RIS-assisted massive MIMO, IEEE Trans. Wirel. Commun. 21 (6) (2021) 4255–4268.

[17] M.A. Cheema, A. Chawla, V.C. Gogineni, P.S. Rossi, Channel estimation in RIS-aided heterogeneous wireless networks via federated learning, IEEE Commun. Lett. (2025) 1, http://dx.doi.org/10.1109/LCOMM.2025.3541873.

[18] H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini, et al., Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design, IEEE Open J. Commun. Soc. 1 (2020) 661–680.

[19] Q. Wu, R. Zhang, Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming, IEEE Trans. Wirel. Commun. 18 (11) (2019) 5394–5409.

[20] W. Chen, C.-K. Wen, X. Li, S. Jin, Adaptive bit partitioning for reconfigurable intelligent surface assisted FDD systems with limited feedback, IEEE Trans. Wirel. Commun. 21 (4) (2021) 2488–2505.

[21] B. Guo, C. Sun, M. Tao, Two-way passive beamforming design for RIS-aided FDD communication systems, in: Proc. IEEE Wireless Commun. Netw. Conf., WCNC, 2021, pp. 1–6.

[22] W. Tang, X. Chen, M.Z. Chen, J.Y. Dai, Y. Han, S. Jin, Q. Cheng, G.Y. Li, T.J. Cui, On channel reciprocity in reconfigurable intelligent surface assisted wireless networks, IEEE Wirel. Commun. 28 (6) (2021) 94–101.

[23] Z. Wang, L. Liu, S. Cui, Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis, IEEE Trans. Wirel. Commun. 19 (10) (2020) 6607–6620.

[24] M.F. Imani, D.R. Smith, P. del Hougne, Perfect absorption in a metasurface-programmable complex scattering enclosure, 2020, arXiv preprint arXiv:2003.01766.

[25] X. Wei, D. Shen, L. Dai, Channel estimation for RIS assisted wireless communications—Part II: An improved solution based on double-structured sparsity, IEEE Commun. Lett. 25 (5) (2021) 1403–1407.

[26] C. Hu, L. Dai, T. Mir, Z. Gao, J. Fang, Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding, IEEE Trans. Veh. Technol. 67 (9) (2018) 8954–8958.

[27] M. Abdussami, R. Amin, P. Saravanan, S. Vollala, BSAPM: BlockChain based secured authentication protocol for large scale WSN with FPGA implementation, Comput. Commun. (2023).

[28] M.T. Hammi, B. Hammi, P. Bellot, A. Serhrouchni, Bubbles of Trust: A decentralized blockchain-based authentication system for IoT, Comput. Secur. 78 (2018) 126–142.

[29] M.A. Cheema, M.K. Shehzad, H.K. Qureshi, S.A. Hassan, H. Jung, A drone-aided blockchain-based smart vehicular network, IEEE Trans. Intell. Transp. Syst. 22 (7) (2020) 4160–4170.

[30] Z. Song, H. Sun, H.H. Yang, X. Wang, Y. Zhang, T.Q. Quek, Reputation-based federated learning for secure wireless networks, IEEE Internet Things J. 9 (2) (2021) 1212–1226.

[31] B. Sun, F. Yu, K. Wu, Y. Xiao, V. Leung, Enhancing security using mobility-based anomaly detection in cellular mobile networks, IEEE Trans. Veh. Technol. 55 (4) (2006) 1385–1396, http://dx.doi.org/10.1109/TVT.2006.874579.

[32] H. Yu, Z. Shen, C. Miao, C. Leung, D. Niyato, A survey of trust and reputation management systems in wireless communications, Proc. IEEE 98 (10) (2010) 1755–1772, http://dx.doi.org/10.1109/JPROC.2010.2059690.

[33] B. McMahan, E. Moore, D. Ramage, S. Hampson, B.A. y Arcas, Communication-efficient learning of deep networks from decentralized data, in: Proc. Int. Conf. Artif. Intell. Statist, 2017, pp. 1273–1282.