

Over-the-air Aggregation-based Federated Learning for Technology Recognition in Multi-RAT Networks

Merkebu Girmay ¹, Mohamed Seif ², Vasilis Maglogiannis ², Dries Naudts ², adnan shahid ², H. Vincent Poor ², and ingrid moerman ²

¹IDLab

²Affiliation not available

December 7, 2023

Abstract

With the continuous evolution of wireless communication and the explosive growth in data traffic, decentralized spectrum sensing has become essential for the optimal utilization of wireless resources. In this direction, we propose an over-the-air aggregation-based Federated Learning (FL) for a technology recognition model that can identify signals from multiple Radio Access Technologies (RATs), including Wi-Fi, Long Term Evolution (LTE), 5G New Radio (NR), Cellular Vehicle-to-Everything PC5 (C-V2X PC5), and Intelligent Transport Systems G5 (ITS-G5). In the proposed FL-based technology recognition framework, we consider edge network elements as clients to train local models and a central server to create the global model. In each client, a Convolutional Neural Network (CNN)-based model is trained from Inphase and Quadrature (IQ) samples collected from a certain combination of RATs. The possible combination of RATs considered in the clients is selected based on the capabilities of the real-world network elements that can be used as a client.

The FL framework involves a process where multiple clients periodically send updates derived from local data to a central server, which then integrates these contributions to enhance a shared global model. This method ensures that the system stays current with the evolving real-world environment while also minimizing bandwidth required for training data transfer and allowing for the maintenance of personalized local models on each client's end.

Over-the-air Aggregation-based Federated Learning for Technology Recognition in Multi-RAT Networks

Merkebu Girmay* Mohamed Seif† Vasilis Maglogiannis* Dries Naudts*
Adnan Shahid* H. Vincent Poor† Ingrid Moerman*

* IDLab, Department of Information Technology, Ghent University – imec, Belgium

† Department of Electrical and Computer Engineering, Princeton University, USA

Corresponding author email: merkebutekaw.girmay@ugent.be (M. G.)

Abstract—With the continuous evolution of wireless communication and the explosive growth in data traffic, decentralized spectrum sensing has become essential for the optimal utilization of wireless resources. In this direction, we propose an over-the-air aggregation-based Federated Learning (FL) for a technology recognition model that can identify signals from multiple Radio Access Technologies (RATs), including Wi-Fi, Long Term Evolution (LTE), 5G New Radio (NR), Cellular Vehicle-to-Everything PC5 (C-V2X PC5), and Intelligent Transport Systems G5 (ITS-G5). In the proposed FL-based technology recognition framework, we consider edge network elements as clients to train local models and a central server to create the global model. In each client, a Convolutional Neural Network (CNN)-based model is trained from Inphase and Quadrature (IQ) samples collected from a certain combination of RATs. The possible combination of RATs considered in the clients is selected based on the capabilities of the real-world network elements that can be used as a client. The FL framework involves a process where multiple clients periodically send updates derived from local data to a central server, which then integrates these contributions to enhance a shared global model. This method ensures that the system stays current with the evolving real-world environment while also minimizing bandwidth required for training data transfer and allowing for the maintenance of personalized local models on each client’s end.

Index Terms—Federated Learning, multi-RAT, Technology Recognition, Spectrum Sensing.

I. INTRODUCTION

As wireless communication continues to evolve, the demand for wireless connectivity and data traffic has been growing dramatically [1]. This surge in traffic, coupled with the finite nature of the available spectrum, poses significant challenges for meeting the increasing demand. To address this issue, spectrum sharing has emerged as a viable solution to optimize spectrum utilization and accommodate the ever-growing traffic. Spectrum sharing enables multiple Radio Access Technologies (RATs) to share the same spectral bands, leading to more efficient use of available resources.

Effective spectrum-sharing decisions can be made based on accurate spectrum-sensing mechanisms. Technology recognition is a spectrum-sensing mechanism used to identify signal types from different RATs. Technology recognition systems can be classified into conventional and advanced machine learning-based solutions [2]. Conventional technology recognition solutions include less complex systems that use mecha-

nisms such as energy detection, cyclostationary feature detection, and matched filter detection to detect unique features of a signal. Despite their simplicity, these conventional mechanisms lead to poor classification accuracy when they are used to identify multiple wireless signal types that do not have easily distinguishable features. For this reason, there has been a recent surge to strengthen the performance of wireless signal recognition using advanced machine-learning-based systems that employ the powerful capabilities of Deep Neural Network (DNN) models [3].

DNN-based wireless technology recognition models follow a series of steps to identify wireless signals [4]. First, a dataset is collected by capturing wireless signals from the devices connected to the considered RATs. Next, the data is pre-processed, and a model is trained to extract key features from these signals, enabling it to recognize patterns that distinguish different RATs. These features are derived from the unique physical layer characteristics of each RAT, including modulation schemes, channel bandwidth, subcarrier spacing, and transmission power. After training and validation, the model is used to analyze incoming signals in real-time, making predictions about the technology in use [5].

Recently, many DNN-based technology recognition solutions have been proposed to recognize signal types from different wireless technologies. For example, a technology recognition model can be used to identify LTE and Wi-Fi signals by continuously capturing Inphase and Quadrature (IQ) samples [6] and using it as an input to a DNN-based classifier.

State-of-the-art DNN-based technology recognition approaches have certain limitations that affect their effectiveness in real-world scenarios. The main limitation of state-of-the-art solutions is their reliance on centralized training approaches [7]. In these approaches, the training process is typically conducted on a central server. However, training technology recognition models necessitates data collection from devices in diverse locations and channel conditions. Hence, centralized training methods involve transferring substantial amounts of data from various sources situated at different locations to a central server. This may result in adding a significant load to the network and possibly impacting its performance. For this reason, most state-of-the-art technology recognition models rely on static datasets that are collected once from simulations

or controlled experimental setups. However, using a static dataset often fails to capture the diversity of real-world network environments, leading to suboptimal performance in practical scenarios. Furthermore, using centralized training places a heavy burden on the server's processing power and storage capacity, potentially leading to scalability issues and increased infrastructure costs.

Federated Learning (FL) has emerged as a promising solution to address the limitations associated with centralized learning [7]. Instead of sending raw data to a central server, FL allows each device to train a local model on its data and only share model updates (gradients) with the central server. These updates are aggregated to improve the global model, which is then redistributed to the devices. Its ability to facilitate distributed learning makes it a suitable choice for numerous wireless communication problems. FL allows the training of technology recognition models on distributed data sources, enhancing robustness and generalization capabilities by incorporating diverse scenarios, channel conditions, and user behaviors. Moreover, the distributed nature of FL makes it more scalable, as the training process occurs locally on client devices, utilizing their computational resources.

To advance in this direction, we propose employing FL for technology recognition in multi-RAT networks. A novel FL-based technology recognition model is developed to recognize signal types from multiple RATs, including Wi-Fi, Long Term Evolution (LTE), 5G New Radio (NR), Cellular Vehicle-to-everything PC5 (C-V2X PC5), and Intelligent Transport Systems G5 (ITS-G5) technologies. For the local models, a Convolutional Neural Network (CNN)-based feature extraction is used to effectively classify the wireless signals. To represent the frequency domain features of each wireless technology, the Fast Fourier Transform (FFT) of the IQ samples captured from each wireless technology is used as an input for the model. Employing a typical Independent and Identically Distributed (IID) FL framework for technology recognition in multi-RAT networks needs a complex receiver at each client to capture, identify, and label samples from each technology. However, our main goal is to create a technology recognition model that does not require such a complex receiver in the clients. Instead, we aim to use a machine learning model-based spectrum sensing technique that can identify the signal type without the need to decode the signal of different wireless technologies from every co-located transmitter.

For this reason, we propose a non-Independent and Identically Distributed (non-IID) FL approach for technology recognition. In the proposed model, each client collects IQ samples from specific combinations of classes. These combinations of classes are determined based on real-world devices used by the considered RATs. For instance, the latest smartphones can be used to capture and label Wi-Fi, LTE, and NR signals. Similarly, an Intelligent Transportation System (ITS) application device can be used to capture and label datasets for C-V2X PC5 and ITS-G5. This approach simplifies the data collection process and improves the feasibility of FL-based technology recognition.

The contributions of this work can be outlined as follows:

- Introducing an approach for distributing a non-IID dataset among clients, taking into account the attributes of actual edge network devices used as clients, and proposing an over-the-air aggregation-based FL framework for a technology recognition model in multi-RAT networks.
- Designing a CNN model structure for local technology recognition models. Local models at the clients are trained based on the FFT of IQ samples captured from signals from a combination of Wi-Fi, LTE, 5G NR, C-V2X PC5, and ITS-G5 technologies, while the central model at the server is obtained by aggregating the local parameters.
- Analyzing the performance of the FL-based technology recognition model using over-the-air aggregation and considering a different number of clients in each aggregation group, a different number of classes in each client, different channel conditions, and different modulation schemes for the aggregation.
- Comparing the accuracy performance and required data transfer requirements of the proposed FL framework with centralized training.

II. RELATED WORKS

FL has emerged as a promising approach for various aspects of wireless communication, including wireless technology recognition and cognitive spectrum sensing [8]. Several studies have explored the potential of FL in various wireless communication scenarios. For instance, in [9], a distributed learning approach is proposed to detect the presence of signal from the Primary User (PU) and estimate the available spectrum in a sensor network, improving overall accuracy while minimizing data exchange between sensors. The results of the study demonstrate that the proposed FL-based approach outperforms separate and autonomous models, leading to a higher accuracy score in terms of spectrum occupancy detection, and data exchange is showcased through the validation dataset.

Similarly, the authors in [7] explore FL in spectrum sensing for cognitive radio environments, highlighting its reliability and privacy preservation. An FL approach is proposed to detect the presence of a 5G downlink signal from a PU and estimate the available spectrum. A similar FL-based PU signal detection scheme is also proposed in [10], [11]. These works introduce FL for cooperative spectrum sensing, reducing the transmission burden while ensuring accurate PU signal detection.

Additionally, in [12], a federated adaptive modulation classification method that preserves privacy with minimal performance loss is proposed, offering a secure alternative to conventional centralized approaches. The signal recognition model is trained to identify four modulation types under class imbalance and varying noise conditions. The results in this paper demonstrate that the solution achieves an average performance loss of less than 2% compared to the conventional centralized modulation classification, making it a promising solution for privacy-preserving adaptive modulation classification.

TABLE I: Related work on FL for wireless signal detection showing a) target detection, b) over-the-air aggregation, c) class imbalance considered, and d) applicability on multi-RAT

Reference	Target recognition	Wireless aggregation	Class imbalance	Multi-RAT
[7]	PU signal detection	×	×	×
[9]	PU signal detection	×	×	×
[10]	PU signal detection	×	×	×
[11]	PU signal detection	×	×	×
[12]	Modulation schemes	×	✓	×
[13]	Modulation schemes	×	×	×
[14]	Image classification	✓	×	×
[15]	Image classification	✓	×	×
Proposed	Wi-Fi, LTE, 5G NR C-V2X PC5, ITS-G5	✓	✓	✓

The study in [13] thoroughly investigates FL’s applicability in signal modulation recognition. The study’s distributed learning approach effectively enhances the accuracy of signal modulation recognition by leveraging an FL-based algorithm. The signal recognition model is trained to identify eleven modulation schemes. Each sample within the dataset consisted of two channels (IQ) of raw data, each with a size of 2×128 . The study conducted simulations to analyze signal modulation identification across various scenarios.

Table I shows the state-of-the-art FL-based signal-type recognition solutions that focus on the identification of the presence of PU signal and the identification of modulation schemes. Despite the use of wireless edge devices as clients, these solutions do not provide over-the-air aggregation mechanisms for the proposed FL systems. In contrast, studies in [14], [15] present FL solutions with over-the-air aggregation. However, these solutions focus on image classification applications for handwritten character recognition and cannot be directly applied to wireless technology recognition.

In this work, we propose a novel over-the-air aggregation-based FL scheme for a technology recognition model developed to recognize signal types from multiple RATs, including Wi-Fi, LTE, NR, C-V2X PC5, and ITS-G5 technologies. We use a non-IID dataset with a class imbalance among the edge clients and varying noise levels are considered to show the performance of the proposed scheme.

III. PROBLEM STATEMENT AND SYSTEM MODEL

Problem Statement: Centralized training in technology recognition leads to challenges like extensive data transfers, compromised generalization, and increased complexity in continuous adaptation to emerging technologies. As a solution to this, the main goal of this paper is to develop an FL-based technology recognition system that can identify wireless signals from multiple RATs using over-the-air gradient aggregation.

In addition, we investigate the impact of wireless channel conditions, the number of clients in the aggregation group, and the number of classes per client on classification accuracy and convergence.

System Model: Fig. 1 shows the framework of the proposed FL-based technology recognition. The figure illustrates that network elements at the edge are represented as clients and are employed to train local models. These clients are commercial user devices that can be connected to Wi-Fi, LTE, NR, ITS-G5, and C-V2X PC5 RATs. Additionally, a central server communicates with all clients and is used for model aggregation. Each client k has a local private dataset $\hat{\mathcal{D}}_k$ with D_k data points, which are obtained by capturing IQ samples. In each round, each client preprocesses the private dataset, uses it to train a local model, and sends the model parameters to the central server. The central server aggregates the model parameters from all clients. The details of the data acquisition, data pre-processing, local model structure, and the proposed federated learning process are presented in Section IV.

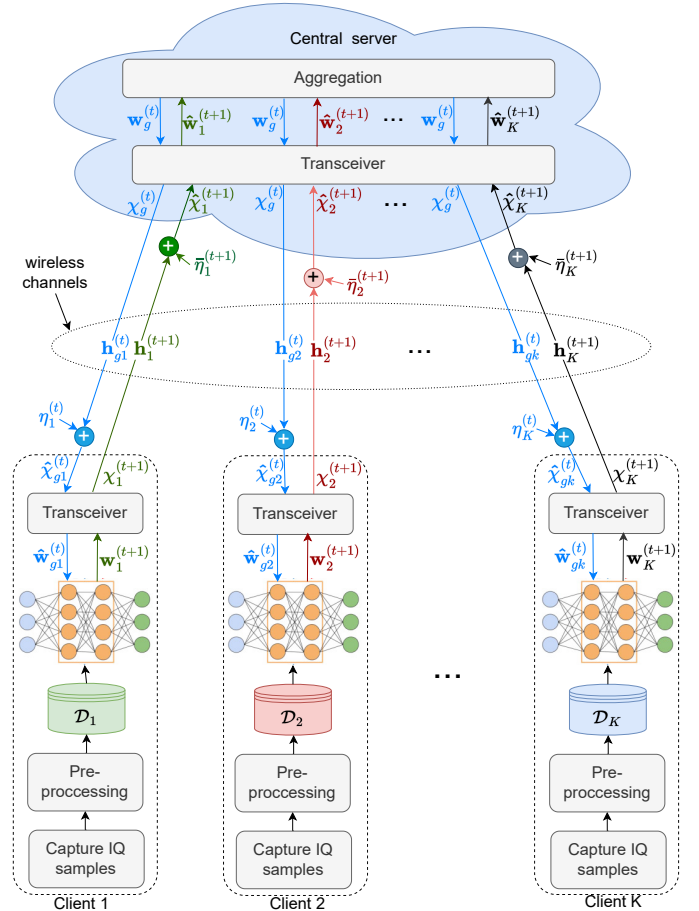


Fig. 1: Framework of the proposed FL scheme: The parameter server and clients exchange the model updates over the wireless channel.

Downlink Transmission: As shown in Fig. 1, in each communication round t , the central server broadcasts the aggregated global parameter vector $w_g^{(t)}$ to all clients. We

consider a wireless FL system where the model parameter exchange between the server and clients is done over a wireless channel. The server modulates the global model parameters $\mathbf{w}_g^{(t)}$ and transmits signal $\chi_g^{(t)}$. Hence, each client k receives a noisy version of $\chi_g^{(t)}$ which can be given as:

$$\hat{\chi}_{gk}^{(t)} = \mathbf{h}_{gk}^{(t)} \otimes \chi_g^{(t)} + \eta_k^{(t)}, \quad (1)$$

where $\mathbf{h}_{gk}^{(t)} = |\mathbf{h}_{gk}^{(t)}|e^{j\phi_{gk}t}$ is the channel coefficient for the downlink channel used to transmit the global parameters to client k and $\eta_k^{(t)}$ is the noise introduced in the downlink channel.

Each client k then demodulates the received signal $\hat{\chi}_{gk}^{(t)}$ and estimates the model parameters. The estimated model parameters at client k are represented by $\hat{\mathbf{w}}_{gk}^{(t)}$ and include errors introduced in communications due to channel noise, imperfect channel estimation, and detection or estimation error. The estimated global model parameter received at client k can be given as:

$$\hat{\mathbf{w}}_{gk}^{(t)} = \mathbf{w}_g^{(t)} + \varepsilon_k^{(t)}, \quad (2)$$

where $\varepsilon_k^{(t)}$ is the effective error vector introduced in the downlink communication on the model parameters received by client k and communication round t .

Uplink Transmission: In the next communication round, each client k updates its local model parameters $\mathbf{w}_k^{(t+1)}$ and sends them to the server. For transmission, the client modulates the local model parameters and generates $\chi_k^{(t+1)}$. The server receives a noisy version of this signal, which can be given as:

$$\hat{\chi}_k^{(t+1)} = \mathbf{h}_k^{(t+1)} \otimes \chi_k^{(t+1)} + \bar{\eta}_k^{(t+1)}, \quad (3)$$

where $\mathbf{h}_k^{(t+1)} = |\mathbf{h}_k^{(t+1)}|e^{j\phi_k(t+1)}$ is the channel coefficient for the uplink channel used to send local parameters from client k to the server, and $\bar{\eta}_k^{(t+1)}$ is the noise introduced in the uplink channel at communication round $t + 1$.

The server uses this received signal to estimate the model parameters sent from each client k , represented as $\hat{\mathbf{w}}_k^{(t+1)}$. The estimated model parameters include errors introduced in the uplink communications. At the server end, the estimated local model parameters received from client k can be given as:

$$\hat{\mathbf{w}}_k^{(t+1)} = \mathbf{w}_k^{(t+1)} + \bar{\varepsilon}_k^{(t+1)}, \quad (4)$$

where $\bar{\varepsilon}_k^{(t+1)}$ is the effective error vector introduced during the uplink communication of the local model parameters transmitted from client k (at communication round $t + 1$).

IV. PROPOSED FEDERATED LEARNING PROCESS

Data Acquisition: We consider real-world network edge devices used as clients. Hence, each client captures IQ samples from signals transmitted from different RATs, identifies the signal type, and labels it. For practical reasons, we consider non-IID FL, where the dataset of each client belongs to a certain combination of classes based on the RATs that can be decoded/identified by different real-world edge devices. As

an example, a smartphone can capture, identify, and label Wi-Fi, LTE, and NR signals. The considered possible combination of classes is presented in Section V.

If we assume K clients are involved in the FL process, each client k captures IQ samples stored as local private raw dataset $\hat{\mathcal{D}}_k$ with D_k signal windows. Throughout the article, note that a signal window refers to a set of M IQ samples captured in a given Time Resolution Window (TRW) and the corresponding label. Hence, a signal window can be denoted as $\hat{\mathcal{D}}_k = \{(\hat{\mathbf{u}}_i^{(k)}, v_i^{(k)})\}_{i=1}^{D_k}$, where $\hat{\mathbf{u}}_i^{(k)}$ is the i -th signal window of client k and $v_i^{(k)}$ is the corresponding label.

The number of classes in a local private raw dataset $\hat{\mathcal{D}}_k$ of each client is defined based on the RATs that can be identified by the client. As each signal window $\hat{\mathbf{u}}_i^{(k)}$ within $\hat{\mathcal{D}}_k$ stores the M IQ samples captured in a duration of the TRW used, the value of M is determined based on the sampling rate and TRW used.

Data Pre-processing: During the pre-processing phase, an M point FFT of the M IQ samples representing each signal window is computed. Hence, the final pre-processed private dataset of each client k used for training can be represented with \mathcal{D}_k which is composed of D_k pre-processed signal windows. The i -th pre-processed signal window, represented as $\mathbf{u}_i^{(k)}$, is obtained using M point FFT of the M IQ samples stored in corresponding signal window before pre-processing window, i.e., $\hat{\mathbf{u}}_i^{(k)}$. Therefore, the l -th value of the $\mathbf{u}_i^{(k)}$ is obtained using:

$$\mathbf{u}_i^{(k)}[l] = \sum_{p=n}^M (I[n] + jQ[n])W_M^{(n-1)(l-1)}, \quad l = 1, 2, \dots, M, \quad (5)$$

where $I[n]$ and $Q[n]$ represent the in-phase and quadrature components of the M IQ samples stored $\hat{\mathbf{u}}_i^{(k)}$ in the data acquisition stage for the i^{th} signal window of client k , respectively, and the term W_M is a complex exponential which can be given as $W_M = e^{(2\pi j)/M}$.

Local ML Model Architecture: A CNN structure designed to efficiently extract features from the FFT of IQ samples representing each signal window is used as a local model. The CNN model comprises three Convolutional layers. In the first layer, there are 64 nested filters, each with dimensions of 1×3 . The second CNN layer consists of 32 nested filters, each with dimensions of 2×3 . The third CNN layer contains 16 nested filters, also with dimensions of 2×3 . These layers are designed to refine the feature extraction process. Between each CNN layer, the Max Pooling layers are used to downsample feature maps, retaining essential information while reducing complexity. The classification phase follows with two Fully Connected (FC) layers. The first FC layer uses ReLU activation, batch normalization, dropout, and regularizer while the second FC layer uses a softmax classifier for probability estimation. Fig. 2 shows how the CNN-based local technology recognition model in each client is used to classify the signal type.

FL Scheme: In each communication round, the FL scheme starts by initializing and broadcasting the global parameters.

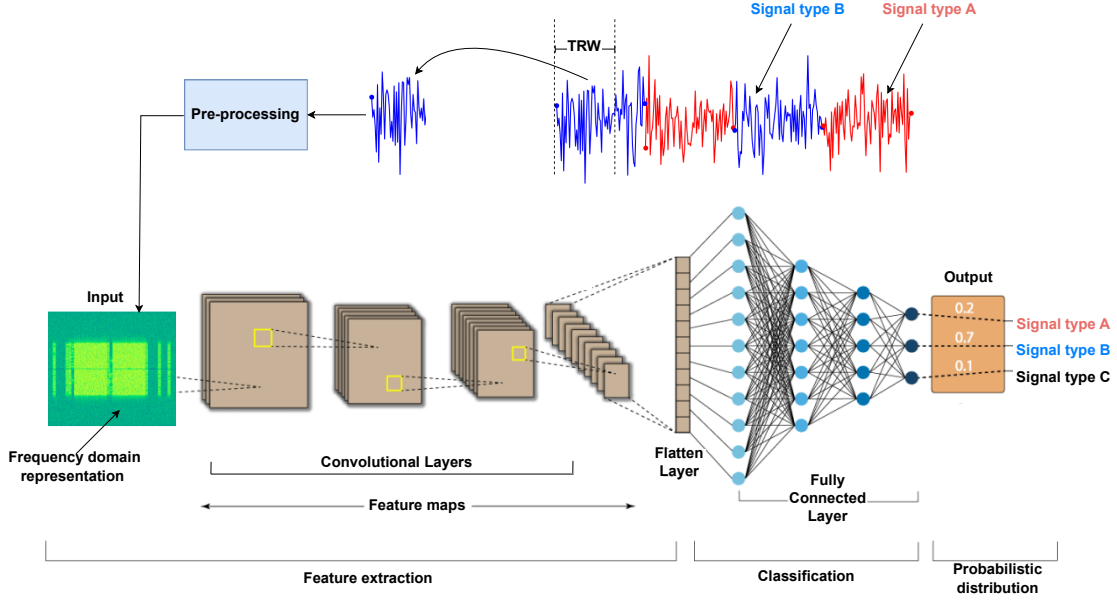


Fig. 2: Technology recognition execution process using the local model in the clients: Each client captures IQ samples in a time resolution window, and the IQ samples are pre-processed using FFT. The CNN-based technology recognition model is used to extract features and classify the signal type.

Algorithm 1: Federated learning for technology recognition in multi-RAT network

Data: Number of communication rounds (T)

Result: Optimal global model \mathbf{w}^*

- 1 Initialize global model $\mathbf{w}_g^{(0)}$
 - 2 **for** $t = 0$ to $T-1$ **do**
 - 3 Broadcast $\mathbf{w}_g^{(t)}$
 - 4 **for** $k = 1$ to K **do**
 - 5 Capture IQ samples for D_k TRWs to get
 $\hat{\mathcal{D}}_k = \{(\hat{\mathbf{u}}_j^{(k)}, v_j^{(k)})\}_{j=1}^{D_k}$
 - 6 Use FFT to get $\mathcal{D}_k = \{(\mathbf{u}_j^{(k)}, v_j^{(k)})\}_{j=1}^{D_k}$
 - 7 Sample a mini-batch $\mathcal{B}_k^{(t)} \subseteq \mathcal{D}_k$ with size n_k
 - 8 Client receives global parameters $\hat{\mathbf{w}}_{gk}^{(t)}$
 - 9 Compute local gradient:
 $\mathbf{g}_k^{(t)} = \frac{1}{n_k} \sum_{j \in \mathcal{B}_k^{(t)}} \nabla f_k(\hat{\mathbf{w}}_{gk}^{(t)}(\mathbf{u}_j^{(k)}, v_j^{(k)}))$
 - 10 Update local model: $\mathbf{w}_k^{(t+1)} = \mathbf{w}_k^{(t)} - \eta_t \mathbf{g}_k^{(t)}$
 - 11 Transmit local parameters $\mathbf{w}_k^{(t+1)}$
 - 12 Server receives local parameters $\{\hat{\mathbf{w}}_k^{(t+1)}\}_{k=1}^K$
 - 13 Aggregate local model updates:
 $\mathbf{w}_g^{(t+1)} = \frac{1}{K} \sum_{k=1}^K \hat{\mathbf{w}}_k^{(t+1)}$
 - 14 $\mathbf{w}^* = \mathbf{w}_g^{(T)}$
-

After that, each client captures IQ samples for D_k TRWs. Then, the captured IQ samples are pre-processed using FFT to obtain the local pre-processed dataset \mathcal{D}_k . Subsequently, the local training is done based on the local datasets and the local

CNN-based technology recognition model, and the local loss function $f_k(\mathbf{w})$ at client k is given by,

$$f_k(\mathbf{w}) = \frac{1}{D_k} \sum_{i=1}^{D_k} f(\mathbf{w}; (\mathbf{u}_i^{(k)}, v_i^{(k)})). \quad (6)$$

where $\mathbf{w} \in \mathbb{R}^d$ is the parameter vector to be optimized. The objective of the FL system is to obtain the optimal global model \mathbf{w}^* at the server by minimizing the global loss function.

$$F(\mathbf{w}) \triangleq \frac{1}{\sum_{i=1}^K D_i} \sum_{k=1}^K D_k f_k(\mathbf{w}). \quad (7)$$

The minimization of $F(\mathbf{w})$ is carried out iteratively. More specifically, in the t -th communication round, the server broadcasts the global parameter vector $\mathbf{w}_g^{(t)}$ over a wireless channel, and each client K receives a distorted $\hat{\mathbf{w}}_{gk}^{(t)}$. User k then computes its local gradient $\mathbf{g}_k^{(t)}$ using stochastic mini batch $\mathcal{B}_k^{(t)} \subseteq \mathcal{D}_k$, with size n_k , i.e., for all $t \in \{1, 2, \dots, T\}$, $|\mathcal{B}_k^{(t)}| = n_k$. Therefore, for all $k \in \{1, 2, \dots, K\}$ and all $t \in \{1, 2, \dots, T\}$, the local gradient $\mathbf{g}_k^{(t)}$ is given by:

$$\mathbf{g}_k^{(t)} = \frac{1}{n_k} \sum_{j \in \mathcal{B}_k^{(t)}} \nabla f_k(\hat{\mathbf{w}}_{gk}^{(t)}; (\mathbf{u}_j^{(k)}, v_j^{(k)})). \quad (8)$$

Subsequently, the local parameter at client k , $\mathbf{w}_k^{(t)} = [w_{k,1}^{(t)}, w_{k,2}^{(t)}, \dots, w_{k,d}^{(t)}]$ is updated according to the update rule $\mathbf{w}_k^{(t+1)} = \mathbf{w}_k^{(t)} - \eta_t \mathbf{g}_k^{(t)}$, where η_t is the learning rate of the distributed learning algorithm at iteration t . This updated local model parameter is then sent over the air to the server. Hence, the server receives a distorted model weights $\hat{\mathbf{w}}_k^{(t+1)}$ and

aggregates the local model updates from the clients to obtain a new global model $\mathbf{w}_g^{(t+1)} = \frac{1}{K} \sum_{k=1}^K \hat{\mathbf{w}}_k^{(t+1)}$. This process is repeated until the model converges, leading to the optimal global model \mathbf{w}^* . Algorithm 1 shows the overall process of the model aggregation in the proposed FL-based technology recognition.

V. NUMERICAL EXPERIMENTS

This section presents the performance evaluation of the proposed technology recognition solution. We used an IQ sample dataset of several RATs, including LTE, NR, Wi-Fi, C-V2X PC5, and ITS-G5 [4]. The dataset was collected using a sampling rate of 20 Msps and 44 μ s TRW. To ensure a comprehensive assessment, 80% of the IQ samples are used for training a local model by each client, and the remaining 20% of the IQ samples are used for evaluating the performance of the aggregated model on the server.

TABLE II: Combination of classes considered for local model training in each client for different numbers of clients

Classes per client	Number of clients	Classes in each client
2 classes per client	3	LTE and NR Wi-Fi and LTE C-V2X PC5 and ITS-G5
	5	LTE and NR Wi-Fi and LTE C-V2X PC5 and ITS-G5 LTE and C-V2X PC5 Wi-Fi and ITS-G5
	10	LTE and NR ($\times 2$) Wi-Fi and LTE ($\times 2$) Wi-Fi and NR ($\times 2$) C-V2X PC5 and ITS-G5 LTE and C-V2X PC5 Wi-Fi and ITS-G5 NR and C-V2X PC5
3 classes per client	3	LTE, NR, and Wi-Fi ITS-G5, Wi-Fi, and LTE LTE, NR, and C-V2X PC5
	5	LTE, NR, and Wi-Fi ITS-G5, Wi-Fi, and LTE LTE, NR, and C-V2X PC5 Wi-Fi, LTE, and C-V2X PC5 Wi-Fi, NR, and C-V2X
	10	LTE, NR, and Wi-Fi ($\times 2$) ITS-G5, Wi-Fi, and LTE ($\times 2$) LTE, NR, and C-V2X PC5 ($\times 2$) Wi-Fi, LTE, and C-V2X PC5 ($\times 2$) Wi-Fi, NR, and C-V2X ($\times 2$)

For each training round, the training dataset is partitioned among multiple clients. To mimic real-world scenarios, we considered that each client has a unique cluster of IQ samples from two or three classes. The dataset cluster can belong to a certain combination of classes, as shown in Table II. To represent distinct locations and channel conditions, each client introduces noise (considering SNR between -15 and 30 dB) to the data. After that, the FFT of the IQ samples from each client is used for the local training process. The local model is trained using a learning rate of 0.001, a batch size of 256, and undergoes 5 epochs.

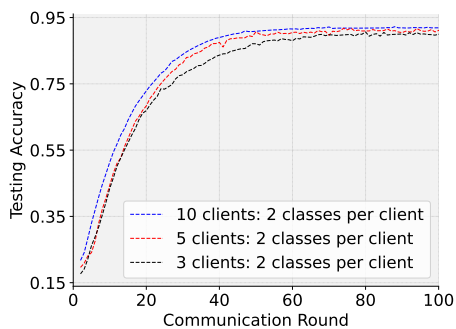
Subsequently, an aggregate model is computed by averaging the local parameters from each client. This process is repeated for each round and enables the proposed FL-based technology recognition solution to deliver robust performance across varying environments. The over-the-air model aggregation process uses wireless model parameters exchange which involves modulation from the sender, introduction of noise at the channel, and demodulation at the receiver side. Considering a wide spectrum of channel conditions, the model parameter communication process involves the introduction of Gaussian noise. This noise is added to the transmission of model parameters, occurring bidirectionally, i.e., based on the channel from the client to the server and reciprocally from the server to the client [16]. This approach is adopted to simulate and account for the unpredictable and often challenging real-world scenarios in which the data transfer takes place.

The training process utilizes a patience of 10 consecutive communication rounds to assess convergence. This is done by monitoring the accuracy of the aggregated model. If the accuracy remains unchanged or fails to improve by at least 0.1%, the process is deemed to have converged.

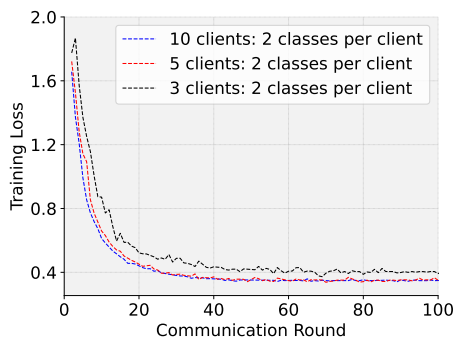
Impact of Number of Clients: Fig. 3 illustrates the loss and accuracy curves for the proposed FL-based technology recognition model for 3, 5, and 10 clients in a group used for computing the aggregated model (for QPSK-based over-the-air aggregation at 10 dB SNR). In this analysis, all the clients are set to have two classes based on the combinations in Table II. The results reveal that increasing the number of clients in the group leads to improved classification accuracy and reduces the number of communication rounds required for convergence. Specifically, the classification accuracy of the optimal aggregated model for 3, 5, and 10 client-based models is 90.27%, 91.61%, and 92.44% respectively. On the other hand, the communication rounds required for convergence are 61, 56, and 43 for 3, 5, and 10 client-based models respectively. This beneficial effect stems from the increased dataset diversity, as each client has its own classes and operates under different channel conditions, thereby contributing valuable information to the model training process.

Impact of the Number of Classes per Client: Fig. 4 presents the accuracy and loss curves for QPSK-based over-the-air aggregation in the proposed FL scheme at a 10 dB SNR wireless channel, considering 2 and 3 classes per client and 5 clients in a group used for computing the aggregated model. The results demonstrate that increasing the number of classes per client significantly improves classification accuracy and reduces the number of communication rounds required for convergence. For instance, the optimal aggregated model for the two classes per client-based clustering achieves a classification accuracy of approximately 91.61% and requires 56 communication rounds to converge. On the other hand, the optimal aggregated model for the three classes per client-based clustering achieves a classification accuracy of approximately 92.74% and requires only 41 communication rounds to converge.

The performance enhancement observed when increasing the



(a)



(b)

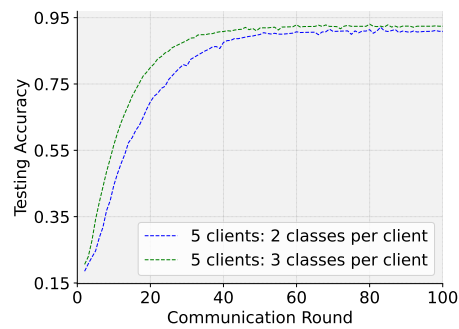
Fig. 3: Impact of the number of clients in a group for over-the-air aggregation-based FL using QPSK modulation at 10 dB SNR: a) testing accuracy b) training loss

number of classes is due to the increased diversity when the number of classes is increased. The increase in the number of classes for each client results in greater data diversity available for training on each client. With a larger number of classes, the client models have access to more diverse data samples, allowing them to learn more generalized representations. The higher number of classes per client compels the client models to learn more complex and abstract features to distinguish between a larger set of classes effectively. As a result, the global model’s performance is boosted as it leverages these informative representations from various clients.

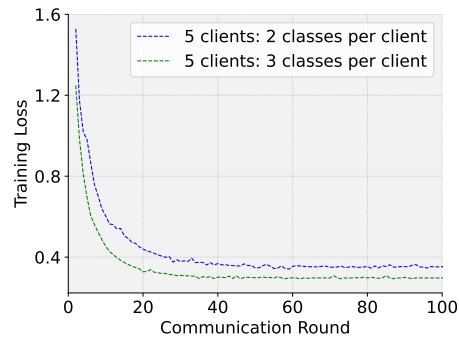
TABLE III: Comparison of QPSK and 16-QAM based over-the-air aggregation for five clients per group and two classes per client

Modulation	Accuracy (in %) for different SNR					
	-5 dB	0 dB	5 dB	10 dB	15 dB	20 dB
QPSK	38.25	43.16	85.27	91.61	91.71	91.84
16-QAM	33.45	34.84	40.28	89.44	91.27	90.68

Impact of Signal Modulation: Table III provides a comparison between the performance of QPSK and 16-QAM modulations used for the over-the-air aggregation. The comparison is based on data obtained using aggregation with five clients, and with each client accommodating two classes. The table presents



(a)



(b)

Fig. 4: Impact of the number of classes per client in FL-based technology recognition for over-the-air aggregation using QPSK modulation at 10 dB SNR channel: a) testing accuracy b) training loss

the accuracy percentages at various SNR levels, ranging from -5 dB to 20 dB. For instance, at 10 dB SNR, QPSK exhibits an accuracy of 91.61%, while 16-QAM achieves 89.44%. This table serves as a valuable reference for evaluating the trade-offs between these two modulation schemes in the context of over-the-air aggregation.

Comparison between Centralized and Proposed FL: We also provide a comparison between the centralized technology recognition model in [4]) and FL-based models, considering scenarios with 3, 5, and 10 clients, each trained with data from 2 distinct classes. For the centralized model, in order to achieve a dataset as extensive and diverse as the one utilized in the case of FL models, we took into account 50 different data collection locations. These locations encompassed a range of SNR, spanning from -15 to 30 dB. Within each of these locations, we utilized 7500 signal windows for every technology (from the dataset used to train the technology recognition model in [4]). Each signal window comprises 1760 values, corresponding to a 20 Msps sampling rate with 44 μ s TRW.

In centralized training, the complete dataset is transmitted to the central server, forming the basis for model training. In the context of the FL-based model, the values are obtained by accumulating the data transfer associated with the parameters transmitted from each individual client to the server, as well as the data exchanged from the server to the clients throughout

all the rounds needed to achieve convergence. For this reason, Table IV shows that training the centralized technology recognition model necessitates 26.42 GB of data transfer, while the FL-based model significantly reduces this burden, with data transfers of only 0.54 GB, 0.75 GB, and 1.05 GB for 3, 5, and 10 clients, respectively. In terms of accuracy, the table demonstrates that the centralized model achieves an accuracy of 94.84%, while the FL-Model (QPSK-based over-the-air aggregation at 10 dB SNR) maintains competitive results with accuracies of 90.27%, 91.61%, and 92.44% for 3, 5, and 10 clients, respectively. Despite utilizing a similar CNN structure in both centralized and FL-based models, the presence of a non-IID dataset distribution among clients results in only a marginal reduction in the classification accuracy of the FL-based model. Generally, the results underscore the potential of FL in minimizing data transfer while preserving high accuracy rates, making it an appealing solution for collaborative machine-learning tasks with distributed data sources.

TABLE IV: Comparison of centralized and FL-based models (QPSK-based over-the-air aggregation for 3, 5, and 10 clients with 2 classes per client)

	Centralized	FL-based model		
		3 Clients	5 Clients	10 Clients
Transferred data	26.42 GB	0.54 GB	0.75 GB	1.05 GB
Accuracy	94.84%	90.27%	91.61%	92.44%

VI. CONCLUSION AND FUTURE WORK

We have proposed a solution to tackle the challenges of centralized training for technology recognition model used to identify wireless signals in multi-RAT wireless networks. In particular, we have considered a non-IID dataset distribution approach on the clients and an over-the-air aggregation-based FL framework, which offers distributed training on client devices to enhance scalability while minimizing data transfer and maintaining classification accuracy. A local CNN-based technology recognition model was trained on the clients based on a dataset from a combination of Wi-Fi, LTE, 5G NR, C-V2X PC5, and ITS-G5 RATs. The possible combination of RATs considered in the data distribution in each client was selected based on the RATs that are used by real-world edge network devices. The evaluation results demonstrate that the proposed FL-based technology recognition greatly reduces the data transfer required for training. The results also show that classification accuracy increases as the number of clients and classes per client increases.

In this work, the performance of an FL-based technology recognition model with over-the-air aggregation is evaluated under different channel conditions. In the near future, this can be extended by proposing adaptive transmission power and modulation schemes for over-the-air aggregation. Furthermore, the bandwidth requirement and classification accuracy analysis of model quantization can be considered as potential future work.

ACKNOWLEDGMENT

This research was funded in part by the Horizon Europe program under the MCSA Staff Exchanges 2021 grant agreement 101086218 (EVOLVE project), by the Belgian Defense through Contract No. 22DEFRA004 (DEFRA-BOLSTER project), and by the U.S. National Science Foundation under Grants CNS-2128448 and ECCS-2335876.

REFERENCES

- [1] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A Speculative Study on 6G," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2020.
- [2] D. Janu, K. Singh, and S. Kumar, "Machine Learning for Cooperative Spectrum Sensing and Sharing: A Survey," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 1, p. e4352, 2022.
- [3] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim, and M. Mokayef, "Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey," *IEEE Access*, vol. 9, pp. 57851–57873, 2021.
- [4] M. Girmay, V. Maglogiannis, D. Naudts, M. Aslam, A. Shahid, and I. Moerman, "Technology Recognition and Traffic Characterization for Wireless Technologies in ITS Band," *Vehicular Communications*, vol. 39, p. 100563, 2023.
- [5] R. K. Saha, "Coexistence of Cellular and IEEE 802.11 Technologies in Unlicensed Spectrum Bands -A Survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1996–2028, 2021.
- [6] M. Girmay, P. Avila-Campos, V. Maglogiannis, D. Naudts, A. Shahid, and I. Moerman, "Intelligent Spectrum Sharing Between LTE and Wi-Fi Networks using Muted MBSFN Subframes," in *Proceedings of the IEEE Wireless and Microwave Technology Conference (WAMICON)*, pp. 13–16, 2023.
- [7] M. Wasilewska, H. Bogucka, and H. V. Poor, "Secure Federated Learning for Cognitive Radio Sensing," *IEEE Communications Magazine*, vol. 61, no. 3, pp. 68–73, 2023.
- [8] S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated Learning for Wireless Communications: Motivation, Opportunities and Challenges," 2020.
- [9] Kułacz and A. Kliks, "Federated Learning-Based Spectrum Occupancy Detection," *Sensors*, vol. 23, no. 14, 2023.
- [10] Z. Chen, Y.-Q. Xu, H. Wang, and D. Guo, "Federated Learning-Based Cooperative Spectrum Sensing in Cognitive Radio," *IEEE Communications Letters*, vol. 26, no. 2, pp. 330–334, 2022.
- [11] T. Yang, W. Zhang, Y. Bo, J. Sun, and C.-X. Wang, "Dynamic Spectrum Sharing Based on Federated Learning and Multi-Agent Actor-Critic Reinforcement Learning," in *Proceedings of the International Wireless Communications and Mobile Computing (IWCMC)*, pp. 947–952, 2023.
- [12] Y. Wang, G. Gui, H. Gacanin, B. Adebisi, H. Sari, and F. Adachi, "Federated Learning for Automatic Modulation Classification Under Class Imbalance and Varying Noise Condition," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 86–96, 2022.
- [13] J. Shi, H. Zhao, M. Wang, and Q. Tian, "Signal Recognition Based On Federated Learning," in *Proceedings of the IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1105–1110, 2020.
- [14] B. Tegin and T. M. Duman, "Federated Learning With Over-the-Air Aggregation Over Time-Varying Channels," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5671–5684, 2023.
- [15] M. S. E. Mohamed, W.-T. Chang, and R. Tandon, "Privacy Amplification for Federated Learning via User Sampling and Wireless Aggregation," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 12, pp. 3821–3835, 2021.
- [16] M. Seif, R. Tandon, and M. Li, "Wireless Federated Learning with Local Differential Privacy," in *Proceedings on the IEEE International Symposium on Information Theory (ISIT)*, pp. 2604–2609, 2020.