FaIR: Federated Incumbent Detection in CBRS Band

*The Department of Electrical Engineering, University of Hawai‘i at Mānoa, Honolulu, HI, 96822, USA

Abstract—The next-generation spectrum access system (SAS) for the Citizens Broadband Radio Service band is equipped with environmental sensors (ESCs) to detect the presence of non-informed incumbent users, which allows the SAS to dynamically reallocate spectrum resource for low privilege users to avoid interference. However, the performance of existing single-node detection model is limited by the sensor’s geo-locations; whereas a naive distributed sensing network with improved detection accuracy introduces a high bandwidth overhead due to the frequent communication of spectrum data. In addition, many existing coherent spectrum sensing methods are not feasible for CBRS band due to the unknown operational characteristics of incumbent military wireless applications.

To address these issues, we propose a machine learning based non-coherent spectrum sensing system: (F)eder(a)ted (I)ncumbent Detection in CB(R)S (FaIR). FaIR leverages a communication-efficient distributed learning framework, federated learning, for ESCs to collaborate and train a data-driven machine learning model for incumbent detection under minimal communication bandwidth. Our preliminary results show that the federated learning method can exploit the spatial diversity of ESCs and obtain an improved detection model comparing to a naive distributed sensing and centralized model framework. We evaluate the FaIR model with a variety of spectrum waveforms at varying SNRs. Our experiments showed that FaIR improves the average detection accuracy compared to the single-node method, using a fraction of the bandwidth compared to the naive multi-node method.

Index Terms—CBRS, Incumbent User Detection, Environmental Sensing Capabilities, Federated Learning

I. INTRODUCTION

To address the increasing demand for connectivity in congested wireless spectrum and to promote dynamic access to spectrum resources, the Federal Communications Commission (FCC) opened the protected spectrum in 3.5 GHz for cellular carriers and opportunistic spectrum sharing. The newly opened spectrum between 3550 MHz to 3700 MHz, also known as Citizens Broadband Radio Service (CBRS) band, was originally occupied by authorized federal and grandfathered Fixed Satellite Service users. Under the rules promulgated by the FCC, these users, now defined as the incumbent access users, will share the spectrum with commercial users and other licensed occupants of CBRS band. However, the incumbent users still have the highest access privilege within the CBRS band and will be protected from harmful interference from users of lower access privileges.

To protect incumbent operations, the FCC prescribes an automated frequency coordinator, known as Spectrum Access System (SAS), to coordinate activities between users in different access tiers. Similar to the case of Television White Spaces management systems, the basic SAS maintains a spectrum geo-database, which registers the locations and operating frequencies of licensed users. The SAS leverages the database to define protection contours for incumbent users. Some incumbent users, such as Navy radar operators, are ship-borne systems with the ability to move from shore to sea. Therefore, the SAS is also equipped with Environmental Sensing Capabilities (ESC) to enable ad-hoc detection of incumbent users.

The ESCs are essentially spectrum sensors that monitor the radio frequency activities within its bandwidth. The sensors identify the unique radio waveform of the incumbent users and report the detected events along with their geo-locations to the SAS, which triggers protective measures. In specific cases, the SAS would require low-tier users to reduce power below the rule limits, force spectrum channel re-assignments, or cease all communications to grant priority to incumbent operations [3]. In July 2019, the FCC approved Federated Wireless, Inc and the collaboration of CommScope and Google to evaluate ESC operations for the 3.5GHz Band [4]. The ESC experimentation of both parties led the FCC to approve the Initial Commercial Deployment (ICD) of the CBRS [5] shortly after.

Spatial diversity of the ESCs is critical in incumbent detection [15]. However, the proposal for the CBRS [3] authorizes a single-node ESC model, which has limited performance when being deployed across a vast region. Under a single-node ESC model, the spectrum datum is collected from one ESC at a fixed geo-location, therefore is prone to errors due to wireless channel impairments such as noise, reflections, fading, and obstructions. Also, a single-node ESC model needs to scan the entire spectrum of the CBRS band, and may miss sporadic signals due to limited sweep speed.

In contrast to the single-node model, a multi-node ESC model employs many sensor nodes to monitor the spectrum simultaneously. Each ESC senses the spectrum from its location and uploads the spectrum data to a central server. The server analyzes the aggregated data to mitigate the errors due to channel fading and improve the detection accuracy. As a result, a naive multi-node ESC model is more accurate in terms of detection but incurs a high communication overhead due to the aggregation of raw spectrum data.

To better addresses the trade-off between accuracy and communication efficiency, we propose a new ESC for the SAS architecture: Federated Incumbent Detection in CBRS (FaIR). We employ a new distributed machine learning technique, Federated Learning, which allows the SAS to obtain a data-driven machine learning model that captures the spectrum characters
of incumbent users, without aggregating the spectrum data from ESC sensors. Our solution distributes the machine learning task to the ESC sensors, allowing the ESCs to collaborate as a loosely coupled federation to train a machine learning model. During training, the Federated Learning procedure enables ESCs to acquire the current detection model from a central server, evaluate it with the local spectrum data, and updates the model parameters intermittently. The central server collects the updates occasionally from the ESCs, computes their average, and applies it to the current model. Throughout the process, raw spectrum data are not uploaded to the central server, which minimizes the communication overhead.

The contributions of the paper are as follows: (1) We compared different spectrum sensing methods, e.g., coherent vs. non-coherent and single-node vs. multi-node, for the task of CBRS incumbent detection, in terms of accuracy, communication overhead, and robustness against unknown waveform.

(2) We proposed a distributed sensing architecture that allows ESCs under geo-spatial constraints to build a machine learning model for incumbent detection. We further provided a learning scheme to minimize the communication overhead on the SAS network, by customizing the federated learning framework.

(3) We evaluated the detection accuracy and communication throughput of the proposed approach through extensive simulations with real-world data captured in laboratory environment. Our experiments showed that FailR improves the average detection accuracy compared to the single-node method, using a fraction of the bandwidth compared to the naive multi-node method.

II. RELATED WORKS

Incumbent detection recently has seen much of its research at federally approved laboratories. Matched filter detection methods, coherent and non-coherent, and support vector machine have been used in [2] and [1] to detect field-measured signals of in-band incumbent radar systems with out-of-band and LTE interference. Both of which use the dataset acquired under the 3.5 GHz Radar Waveform Capture final report [7] [8] provided to Department of Commerce. Sensitivity and placement of ESC sensors has been studied [15] for improving ESC network design to protect protect incumbents given sensor constraints.


Rajendran et al. [21] showed that for high signal-to-noise ratio (SNR) conditions, simple Long Short Term Memory (LSTM) models performed well enough in side by side comparison to complex Convolutional Neural Networks (CNN) LSTM models from [27] and that CNN models performed well in low SNR conditions (below -2dB), as in [17]. The work of Lees et al. [13] present an evaluation report on incumbent SPN-43 Navy radar detection using classical signal detection theory, deep neural networks, and machine learning algorithm using a proprietary spectrogram dataset, rather than I/Q vector data.

Federated learning has been gaining more traction from its recent introduction in McMahan et al. communication efficient learning work [14]. Konecny et al [12] explored reducing the communication costs between highly distributed sensor networks with many distributed client nodes using methods of quantization, random rotations, and sub-sampling.

III. SAS AND INCUMBENT DETECTION

Based on the FCC rules, wireless users in CBRS band must connect to the SAS to allow dynamic spectrum coordination. These users have full access to their allocated CBRS channels if no incumbent operators present within the spectrum. The SAS leverage the ESC for detecting the incumbent waveform presence within the CBRS band and manages spectrum allocation/reallocation for all CBRS devices.

The ESC sensors monitor the spectrum and periodically send the spectrum data to the SAS. Once detecting the presence of incumbent waveform within the spectrum data, the SAS enforces FCC policy by re-assigning the wireless channels, requesting the CBRS devices to back off from the interfering frequency or demanding them to cease transmission as necessary.

However, incumbent detection in CBRS band is difficult, as the majority of incumbent systems are for military purposes and have dynamic or unknown operational characteristics. For instance, NTIA’s Office of Spectrum Management and Institute for Telecommunication Sciences has researched the incumbent radar emissions characteristics for several use cases in ESC deployment, and found that the radar signal characteristics vary from system to system [22]. In reality, federal radar systems are not required to submit parameter details or characteristics of the transmission signal per the conditions provided in [3].

As a result, existing proposals for SAS mostly rely on non-coherent methods, such as energy, pulse width, pulse reception rate, and beam scanning interval [23], which require little prior knowledge of the signals transmitted. However, the detection accuracy of these methods is lower compared to coherent approaches due to the limit of SNR wall [26]. To improve the performance, recent results in machine learning show that non-coherent signal detectors can be augmented by sophisticated data-driven models such as deep neural networks [17]. Unfortunately, the training process to obtain such models often requires a large volume of spectrum data, which cannot be accommodated by existing SAS network infrastructure.

IV. SYSTEM DESIGN GOALS

To improve detection accuracy and lower communication overhead, we propose a distributed CBRS SAS system with non-coherent incumbent detection capability. The method integrates a deep learning model with a distributed federated
learning framework to achieve accurate waveform classification with minimal communication bandwidth. The design goals of the system are listed as the following.

**Non-coherent Detection** An essential feature of our design is the detection method remains non-coherent and requires no prior knowledge of the incumbent signals. Per the FCC conditions provided in [3], federal radar systems operating in CBRS band do not have to provide the SAS with parameter details or characteristics of the transmission signals. This ruling limits the choices of CBRS incumbent detection to non-coherent methods, which are not robust against variations in channel noise and fading processes [26]. However, this disadvantage can be partially addressed through self-adaptive algorithms, such as machine learning, which automatic reconfigure the detector parameters based on observed data [13], [17], [19].

**Distributed Detection** Our design aims to construct a data-driven model of the incumbent signals through a distributed multi-node approach. The detection capability of a single ESC can be affected by wireless channel impairments such as noise, spatial diffusion, fading, interference, noise, or the mobility of incumbent users. An ESC sensor experiencing shadowing or fading effects cannot distinguish between a vacant spectrum and deeply faded incumbent signals. A distributed sensing strategy compensates the wireless channel effects by placing the ESC sensors across the region of interest. The spatial diversity increases the likelihood for sensors to capture clear samples of the incumbent signals, which allows the machine learning algorithm to build an accurate signal model to facilitate detection.

**Low Communication Overhead** Centralized spectrum sensing strategies are taxing to the SAS network since the ESCs need to transfer a large volume of spectrum data to the central SAS server for model training. Based on the database maintained by the Wireless Innovation Forum, many ESCs will be deployed into areas in remote locations, often under extreme topography, congested wireless traffic, or lack of cellular network connections. As a result, the backhaul of the SAS network permits limited bandwidth for data communication, which is not suitable for centralized data aggregation or data-intensive distributed schemes. Federated learning by design requires neither data aggregation nor the participation of every client during model training [14], which significantly reduce the data throughput on the SAS network.

**High Detection Accuracy** Correctly detecting and avoiding the incumbent users is imperative to successful operations in the CBRS band. A fine-tuned detector ensures accurate detection and fast reaction from the SAS, which can be achieved through an iterative algorithm optimizing the model parameters. The use of federated averaging, (discussed in VI-A), provides regularization benefits similar to the dropout techniques [14], which improves the training algorithm’s robustness against data noise and produces more accurate models for incumbent detection.

**V. THE FAIR ARCHITECTURE**

The FaIR is an intelligent SAS architecture, which leverages the computation power of ESC sensors to enable machine learning based incumbent detection. It comprises of three main components as shown in Figure 1: 1) multiple ESC sensors deployed at different locations, e.g., the FaIR Sensor/s (FS/s); 2) a federated SAS server, e.g., the FaIR Master (FM), 3) the network between FSs and FM, e.g., the FaIR network.

1) **FaIR Sensors** (FSs) are distributed across the region with commercial devices operating in the CBRS band. As a distributed sensor, the FS offers a sensing solution to the unique environment of the CBRS devices. The sensors geo-location allows the sensor to learn a unique spectrum environment. The disturbed sensors run both a federated global model and the new localized model to evaluate performance and select the optimal model to deploy.

2) **FaIR Master** (FM) is the centralized server that learns from the distributed FS by using a technique called federated averaging. The server controls the frequency of the federated averaging protocol. The FM also provides each FS with an initial base model to initialize the learning process for each FS.

3) **FaIR Network** facilitates the interaction between the FSs and FM. The FM provides each FS sensor with a base model to apply a technique known as transfer learning, known to benefit a distributed learning environment [6], which initializes the neural network parameters, \( w_0 \). Then, the sensor applies new spectrum data to the model to learn the unique environment. The FS trains its own local neural network until satisfactory performance is evaluated and then the FS is determined eligible to send parameter updates to the FM server to be averaged. The participation in the federated averaging protocol is strictly up to the clients, where not every client participates. The server then averages the weight contributions of participating clients. We discuss the advantage on the SAS communication load using the federated averaging protocol of FaIR in Section VII-E.

**VI. DETECT INCUMBENT SIGNALS THROUGH FAIR**

To detect incumbent signals, the FaIR employs the federated learning techniques to train a classification model to differen-
tiate the waveform captures in CBRS band. The decentralized approach, known as federated learning, coined by McMahan et al. [14], allows training to continue at the client level. Over time, clients contribute new parameter updates to build a centralized model. Federated learning is characterized as a low-communication as clients participate according to a determined criterion. A naive distributed sensing network that generates a heavy traffic load would be taxing to the SAS. Thus, the defined federated learning framework details a communication efficient method to minimize the congestion in the SAS.

A. Model Learning

The objective function of the detection model is to classify a waveform, \( x_i \) as one of the known waveforms of certain modulation scheme, \( y_i \). In particular, our interests is to identify the modulation scheme of the incumbent operator. Therefore, we construct our neural networks to minimize the loss \( l(x_i, y_i; w) \) of each client model due to some prediction with model parameters \( w \) on the waveform and the predicted modulation.

If we have \( C \) distributed ESC clients and \( K \) eligible participants for federated averaging where \( n_k \) defines the sample size for a participant, \( k \in K \). Then we want to minimize the loss function for each sensor, \( k \), such that

\[
\min_{w \in \mathbb{R}^d} J_k(w) \quad \text{where} \quad J_k(w) = \frac{1}{n_k} \sum_{i=1}^{n_k} l(x_i, y_i; w),
\]

The aggregated loss function of the entire system can be defined as \( F(w) = \sum_{k=1}^{K} n_k J_k(w) \).

A common approach in distributed learning, as in [24] and [20], is to take the gradients \( \nabla J_k(w) \), of each participating client, \( \forall k \in K \), training step and average them at the server to update the model \( W^{t+1} = W^t - \eta \nabla F(w) \), for \( \nabla F(w) = \sum_{k=1}^{K} \frac{n_k}{N} \nabla J_k(w) \).

We define \( w_k^t \) as the weight parameters of of the \( k \)-th participant and \( W^t \) as the weight parameters of the aggregated global model at time \( t \). Since a constant communication of gradient updates after every training step would put a strain the SAS, we take advantage of an equivalent form such that for each client update, we have that \( \forall k \in K, w_k^{t+1} = W^t - \nabla J_k(w^t) \). Then we can write the average of all clients weights as \( W^{t+1} = \sum_{k=1}^{K} \frac{n_k}{N} w_k^t \). This allows each client to perform multiple local updates, \( w_k^{t+1} = w_k^t - \eta J_k(w_k^t) \), on local training data before sending updated parameters to the centralized server. This result is called federated averaging and provides communication-efficiency to a distributed learning framework.

B. Signal Classification

When the distributed ESCs are trained, they can leverage the distributed learning framework to detect the incumbent user in new spectrum data. Once the incumbent user is detected, the FaIR network reports to the SAS for spectrum reassignment when detected.

**VII. EXPERIMENT PERFORMANCE AND EVALUATION**

We evaluated the detection accuracy and communication throughput of the FaIR through extensive simulations with real-world data captured in laboratory environment. The performance of the FaIR architecture is evaluated based on the problem of signal modulation detection.

A. Test Environment

Our laboratory setup, as shown in Figure 2 is designed to test and build spectrum sensing environments using waveforms generated by signal generators and software-defined radios (SDRs). We created one incumbent user using a Keysight N5192A signal generator and two lower-tier users using two Ettus N210 USRP. We created two ESC sensors using two additional Ettus N210 USRPs and a Keysight N9010B spectrum analyzer. Through the internal network, this setup allows us to generate the synthetic modulated waveforms, capture the results at the ESC sensors, and analyze fading effects at different positions.

B. Dataset

The public use of CBRS is in its infancy. Currently, there is infrequent occupancy within spectrum, i.e. occasional Navy Radar systems, which are difficult to obtain. Also, there is no publicly available I/Q CBRS data, including datasets with incumbent waveforms. However, we assume (Navy) radar systems to be identified as unique pulse modulated radar or a radar systems that is not modulated, e.g. continuous wave radar (Doppler). With the test environment in Figure 2, we generate synthetic versions of real-world modulated waveforms by transmitting a generated signal over the channel and capturing the signal under incurred channel affects. Our synthetic waveform generation and waveform captures are similar to the work of O’shea et al. [19] using GNU Radio.

The captured waveforms are used in conjunction with the signals in the RML2018.10a [18] dataset, a recent modulation dataset used for waveform modulation learning tasks, to show
displays the benefit of deep convolutional architectures. The Shallow CNN, with the most poor performance, surpassing 70% accuracy at a SNR of 2dB.

The base model is trained on 400,000 randomly sampled data from our lab environment captures and the RML2018.10a dataset. The training set then contains a randomly sampled distribution of different signal to noise ratio spectrum captures for each modulation class. The data consist of signal I/Q vectors where the modulated signals are generated at a 4 samples/symbol and sample length of 1024, i.e. 2x1024 I/Q vector.

C. Deep Learning Models

First, a base model is selected for the FaIR architecture. The base model is the deep learning architecture selected for the distributed network (i.e. ResNets). We explore the use of several SoA deep neural network architectures found in recent literature including shallow convolutional neural network (CNN) [19], deep CNN (in the style of the VGG CNN) [25], and residual network (ResNet) [18]. For brevity, we direct the reader to the references above for the neural network builds.

D. Results

The base model is trained on 400,000 randomly sampled data from our lab environment captures and the RML2018.10a dataset. The training set then contains a randomly sampled distribution of different signal to noise ratio spectrum captures for each modulation class. We select the number of clients, FaIR sensors, to be 10. Each client then runs transfer learning with a unique 60,000 samples. To demonstrate performance of federated learning, clients submit their parameter updates after some time, $T$. Then, the server aggregates the client model weights, takes the federated average, and generates a new global model. The new global model is evaluated with a unique test set, unique to all sensors $k \in K$.

Figure 3 shows the performance of the best performing clients using the VGG, ResNet, and Shallow CNN algorithms at varying SNRs. The best performing architecture is the VGG network, surpassing 70% accuracy at a SNR of 2dB. The ResNet performance is slightly below that of the VGG network. The Shallow CNN, with the most poor performance, displays the benefit of deep convolutional architectures.

We show that federated learning performs as well as current state of the art research. We expect that federated learning can improve performance in deployment, where sensors would be distributed across different locations of a county. We expect from the results of adding more clients to increase the gap between a single client model performance and the federated average model as a result from neural networks behavior given a larger training set. This topic is a topic for future work in this study.

E. Communication Load

To analyze the communication-efficiency of the FaIR, we first look at the communication load of the naive distributed sensing framework. To generate a centralized model in the naive disturbed sensing framework, the sensors collect raw spectrum data and send the data through the ESC network to build a centralized model. The clients are in-sync with the server for real-time model building with $k$ client sensors and $n_k$ samples per client. We determine the communication load over some time interval for the naive distributed sensing framework to be $N \times D_w$, where $N = \sum_{i=1}^{k} n_k$ is the number of waveform samples and $D_w$ is the data size of the waveforms, assumed to be the same for each sample.

Similarly, we can determine communication load for the FaIR architecture and the out-of-sync server updates via the federated averaging protocol. We have the same number of clients $k$, the server requests the federated averaging protocol $M$-many times, and the clients parameter/weight has data size $D_k$. Then the communication load for the FaIR architecture is $(kM) \times D_k$. With a complex deep learning network, generally $D_w < D_k$ however the number of updates $kM << N$. That is, $kM$ is significantly smaller than $N$ and therefore the volume of the naive distributed sensing model is much larger.

Figure 4 demonstrates the improved accuracy of using a federated approach vs selecting the best performing model from the set of participating sensors, i.e. $W^{t+1} = \sum_{k=1}^{K} \frac{w}{n_k} w_k^{t+1}$ vs $W^{t+1} = w_i^{t+1}$, where $w_i^{t+1}$ is selected as the optimal edge sensor from the set of all participating sensors, $i, k \in K$.

Communication efficiency of the FaIR frame work is achieved by the following. First, sending weight parameter...
updates does not congest the SAS by eliminating frequent raw spectrum data communicating to a centralized server. Second, allowing the FS to train on many captured waveforms before averaging rather than averaging at every iteration, detailed in Section VI-A, reduces the volume of communication to the server. Last, not every FS participates in the federated averaging protocol, only those that meet satisfactory conditions.

VIII. CONCLUSION

In this work, we provide a new SAS architecture for communication efficient incumbent detection in CBRS band, FaIR. FaIR augments existing non-coherent spectrum sensing methods with data-driven machine learning model to accurately detect incumbent users with unknown operational parameters. It leverages a distributed sensing framework to collects multiple CBRS spectrum data at different locations to minimize errors due to channel noise and fading effect. Most importantly, FaIR demonstrates the feasibility to use communication-efficient learning protocols, such as federated averaging, to significantly reduce the communication overhead without jeopardizing the incumbent detection accuracy.

Our preliminary results show that state of the art spectrum classification algorithms in a federated learning environment performs very well using our test environment generated and captured waveform of In-phase and Quadrature (I/Q) data and a recent radio classification dataset [18]. Our experiments using a deep learning networks trained on 11 modulated signals at varying SNRs shows improved accuracy from a the optimal single client model to the federated averaged model of 76.92% to 78.63% respectively.

IX. FUTURE WORK

The emerging field of deep learning for spectrum applications has limited dataset resources and, to our knowledge, no dataset that is publicly available for CBRS (3.5-3.7 GHz) learning tasks specifically. Our laboratory testbed allows us to collect and generate real world-data in the CBRS. We plan to collect more waveform data monitored from distributed sensors to analyze the true potential of federated learning in the CBRS. We would also like to investigate the relationship of accuracy and client participation in the federated averaging. We believe there to be a positive relationship between the number of clients and the federated averaged model performance based on our small subset of client participation results.

REFERENCES