Federated Learning for Spatial Reuse in a multi-BSS (Basic Service Set) scenario

Selim F. Yilmaz[1], Emre Ozfatura[1], Kerem Ozfatura[1], Ozlem Yildiz[2] and Deniz Gündüz[1]

[1]Department of Electrical and Electronic Engineering, Imperial College London, UK

[2]Department of Electrical and Computer Engineering, New York University, US

TU AI/ML in 5G Challenge 2021

Motivation and Preliminaries

- Basic service set (BSS): a wireless network where clients connected through an access point AP and referred as station the (STA).
- In dense deployment of WLAN, multiple Basic Service Sets (BSSs) might be located in the overlapping areas.
- Overlapping BSSs (OBSSs): channel access mechanisms based on Carrier Sense Multiple Access (CSMA) underutilizes the spectral resources.
- Spatial reuse operation (SR) included in the IEEE 802.11ax-2021 (11ax) amendment aim to increase the spectral efficiency by allowing parallel transmissions.

SR with OBSS Packet Detect (PD)

- Clear Channel Assessment CS (CCA/CS) protocol aim to prevent any Wi-Fi device to transmit while another one is already transmitting on the same channel.
- CCA/CS threshold: to decode the preamble of another transmitting device and the consider medium as busy.
- OBSS/PD threshold: to ignore a inter-BSS transmission and seek SR opportunity.
- Challenge: how to decide on OBSS/PD threshold

SR with OBSS Packet Detect (PD)

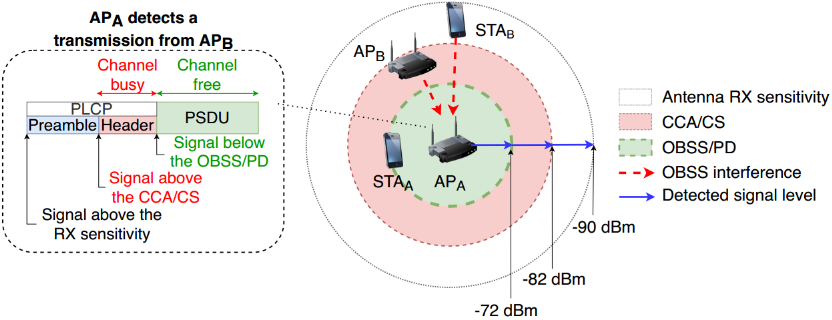

Figure 1: Visualization of the detecting a transmission and looking for SR opportunity using CCA/CS and OBSS/PD thresholds [Wihelmi et. al. 2021].

Performance Analysis

- A given transmission is considered successful if:
  - The power sensed at the receiver from the frame being decoded remains above the CCA/CS.
  - Signal-to-Interference-plus-Noise Ratio (SINR) stays above the capture effect (CE) threshold [Wilhelmi et al. 2019]
- Throughput of the successful transmission are determined according to modulation coding scheme (MCS) based on SINR and RSSI pair.

Federated Learning (FL)

- sample clients $\in \mathcal{S}_t$
- Perform H local iterations

Algorithm 1 Federated Averaging (FedAvg)

for $t = 1, 2, \ldots$ **do**
  for $i \in \mathcal{S}_t$ **do** in parallel
    Pull $\boldsymbol{w}_t$ from central server: $\boldsymbol{w}_{i,t}^0 = \boldsymbol{w}_t$
    for $\tau = 1, \ldots, H$ **do**
      Compute SGD: $\mathbf{g}_{i,t}^{\tau} = \nabla_{\boldsymbol{w}} f_n(\boldsymbol{w}_{i,t}^{\tau-1}, \zeta_{i,\tau})$
      Update model: $\boldsymbol{w}_{i,t}^{\tau} = \boldsymbol{w}_{i,t}^{\tau-1} - \eta_t \mathbf{g}_{i,t}^{\tau}$
    Push $\boldsymbol{w}_{i,t}^H$
  Federated Averaging: $\boldsymbol{w}_{t+1} = \frac{1}{|\mathcal{S}_t|} \sum_{i \in \mathcal{S}_t} \boldsymbol{w}_{i,t}^H$

Problem Definition

- Objective: Finding an OBSS/PD threshold for the BSS of interest that maximize the average throughput of the of STAa for measured RSSI and SINR values.
- Method: using a neural network (NN) architecture to predict throughput of the STAs.
- Approach: utilize the FL framework to train NN.
  - Context: an OBSS scenario with a BSS of interest (containing AP_A).
  - Each context (with corresponding AP_A) considered as user participating FL.
  - The local dataset of each context consist of RSSI and SINR values for the STAs of interest, interference at the AP of interest, set of possible OBSS/PD thresholds and the corresponding throughput values for STAs.

Problem Formulation

- Our objective is to find best threshold configuration that maximize the throughputs:

$$\arg\max_{\boldsymbol{\tau}} \sum_{i=1}^{n} \sum_{j=1}^{s_i} t_{j,\tau_i}^{(i)}, \tag{1}$$

- $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_n]^T$a and τ_i is the threshold of context i.
- s_i is the number of STAs per AP in context i.

Problem Definition - 2

- We cannot directly calculate or know the throughput $t_{j,\tau}^{(i)}$.
- Thus, we estimate it via a model $\hat{t}_{j,\tau}^{(i)} = f_{j,\tau}^{(i)}$
  - i is the context index
  - j is the index of STA connected to $\mathrm{AP_A}$
  - τ is the threshold

Methodology - 1

- Estimating one STA's throughput is highly related to estimating another STA's throughput in the same context.
- Thus, we use mult-output regression to exploit this relation.
- Multi-output regression based formulation:

$$\boldsymbol{f}_{\tau}^{(i)}(\boldsymbol{W}_i) = \left[f_{1,\tau}^{(i)}\ f_{2,\tau}^{(i)} \dots f_{b,\tau}^{(i)}\right]^T, \qquad (2)$$

- i is the context index
- $\boldsymbol{W}_i$ is the neural network weights of i^{th} context

Methodology - 2

Since every context may have different number of STAs per AP, we mask the nonexistent STAs as the following:

$$f_{k,\tau}^{(i)} = \hat{t}_{k,\tau}^{(i)} = t_{k,\tau}^{(i)} = 0, \quad \forall k \in \{s_i + 1, \ldots, b\}.$$

This way, we do not backpropagate any loss for nonexistent STAs and become able to train our model for variable number of STAs per AP for every context.

Methodology - 3

- Then, we define the ground truth vector as:

$$\boldsymbol{t}_{\tau}^{(i)} = \left[t_{1,\tau}^{(i)} \ t_{2,\tau}^{(i)} \ \dots \ t_{b,\tau}^{(i)} \right]^T$$

- For the context i (local node), our objective is to minimize mean-squared error for regression task, i.e.,

$$\underset{\boldsymbol{W}_i}{\arg\min} \left|\left| \boldsymbol{f}_{\tau}^{(i)}(\boldsymbol{W}_i) - \boldsymbol{t}_{\tau}^{(i)} \right|\right|_2^2 .$$

- We use a feed-forward neural network as our model $\boldsymbol{f}_{\tau}^{(i)}(\boldsymbol{W}_i)$ with weights $\boldsymbol{W}_i$.

Simulations - The Dataset and Implementation Details

- For training and validation: Scenario 3 dataset
- 80% training, 10% validation 1, 10% validation 2
- Validation 1: early (global) stopping
- Validation 2: hyperparameter tuning
- Test Dataset: Test scenario
- Evaluation Metric: Mean Absolute Error

Simulations - Competition Results

Team	MAE (Mbps)
FedIPC (Ours)	**5.8572**
FederationS	6.5534
WirelessAI	8.9130

Table 1: ITU-ML5G-PS-004 Competition Results