IEEE P802.11
Wireless LANs

|  |
| --- |
| Proposed IEEE 802.11 AIML SC Technical Report Text for Federated Learning  |
| Date: 2025-07-28 |
| Author(s): |
| Name | Affiliation | Address | Phone | email |
| Rui Yang | InterDigital Inc. |  |  | rui.yang@interdigital.com |
| Alphan Sahin | University of South Carolina |  |  | asahin@mailbox.sc.edu |
| Xiaofei Wang | InterDigital Inc. |  |  | Xiaofei.wang@interdigital.com |

Table of Contents

[Executive Summary 3](#_Toc181693498)

[1 Introduction 4](#_Toc181693499)

[1.1 Terminology 4](#_Toc181693500)

[1.2 Background information 4](#_Toc181693501)

[1.3 Technical Report Overview 4](#_Toc181693502)

[2 AIML use cases for IEEE 802.11 5](#_Toc181693503)

[2.1 General 5](#_Toc181693504)

[2.2 Use case that enables AIML in WLAN 5](#_Toc181693505)

[2.2.1 Use case 1: 5](#_Toc181693510)

[2.3 Use cases that leverage AIML to enhance WLAN performance 5](#_Toc181693511)

[2.3.1 Use case 2: 5](#_Toc181693512)

[3 Conclusions and Recommendations 7](#_Toc181693513)

[3.1 Conclusions 7](#_Toc181693514)

[3.2 Recommendations 7](#_Toc181693515)

[References 8](#_Toc181693516)

# Executive Summary

# Introduction

## Terminology

AIML Artificial Intelligence/Machine Learning

CSI Channel State Information

ED Edge Device

ES Edge Server

FL Federated Learning

DL Downlink

OAC Over-the-Air Computation

UHR Ultra High Reliability

UL Uplink

## Background information

## Technical Report Overview

# AIML use cases for IEEE 802.11

## General

The AIML TIG has identified in its technical report two categories of use cases for WLAN in the area of AIML [1] . In the first category, it is explored how WLAN can efficiently enable and support AIML algorithms that need data across the whole or part of the network and make AIML models available to the IEEE 802.11 devices within the network. One use case has been identified for this category. In the second category, it is explored how AIML-based methods may be leveraged to improve the performance of various operations or layers in WLANs. Four use cases have been identified for this category. Additional use cases for AIML in these categories and additional study results for the use cases described in [1] are described below.

***Editor: The paragraphs below are proposed to be added to the AIML SC technical report.***

## Use case that enables AIML over WLAN

1.
2. 1.
	2.

### Use case 1: Federated Learning over WLAN

#### Use case description

Machine learning (ML) has recently been viewed as one of the most essential technologies for many emerging applications, including autonomous driving, augmented reality, wireless sensing, and general-purpose robotics. Using ML in these applications may require high throughput and/or low latency wireless connectivity between data-generating nodes (e.g., IoT devices, virtual reality headsets, sensors, smartphones) and processing units (e.g., servers, routers), especially when the underlying computations (e.g., training a neural network over a wireless network) are significantly heavier than those required by most current applications.

To solve this problem, as a general theme, distributed optimization over large-scale networks with paradigms, such as federated learning (FL), has been heavily studied in academia and industry [1-3]. Various strategies have also been investigated in industrial standards, including IEEE and 3GPP [4-5], but currently only at the network architecture level or as use cases to assess the system requirements. To enable those new paradigms over standardized wireless communication networks, the existing radio technologies in standards, such as Wi-Fi, including PHY and MAC layer protocols, need to be enhanced or redesigned to overcome some challenges introduced by AI/ML enabled applications. Hence, some advanced PHY and MAC protocols based on fundamental wireless technologies that support ML paradigms may need to be introduced into those standards. To this end, there is growing attention to efficient computation methods over wireless technologies [6-7].

Traditionally, communication and computation are viewed as separate tasks. This approach has been very effective from the engineering perspective, as isolated optimizations can be performed. However, for many computation-oriented applications, the ultimate goal of the network is a mathematical function (e.g., arithmetic mean, maximum, minimum) of the local information distributed at the different devices rather than the local information itself. In such scenarios, information-theoretical results show that intended “interference” in a multiple-access channel for computation, i.e. over-the-air computation (OAC), can provide a significantly higher achievable computation rate than separating communication and computation tasks [8]. With this motivation, the rest of this subclause discusses feasibility to enable OAC in wireless networks for FL, particularly over WLAN and some practical issues [9].

FL is a type of distributed learning paradigm, where a neural network is aimed to be trained by using the local data generated at Edge Devices (EDs) without moving them to a centralized server. With FL, instead of datasets, model parameters or gradients are transmitted back and forth between EDs and an Edge Server (ES). Since local data never leaves where it is generated, FL promotes data privacy for EDs.

Federated Learning based on stochastic gradient descent (FedSGD) can be explained as follows:

* Step 1: ES, e.g., an AP, broadcasts the model parameters
* Step 2: Each ED, e.g., a STA, receives the model parameters and calculates the gradients based on its local dataset
* Step 3: Each ED transmits the gradients in the uplink direction to the ES
* Step 4: ES aggregates the received gradients (e.g., arithmetic mean or weighted average) and updates the model

The abovementioned procedure continues until a convergence criterion is met (e.g., the loss function is less than a threshold). Steps 1-4 are illustrated in Figure 1.



Figure 1. Steps for federated learning over a wireless network based on gradient aggregation

The rationale behind the FedSGD can be explained as follows:

Let and be the local dataset at the th STA and the global dataset, i.e., the union of local dataset, i.e., . Let be the data sample (e.g., an image) and the corresponding label (e.g., chair) in . In a centralized setting where all data is available at the ES, a neural network training problem can be expressed as a general optimization problem given by

where is function that measures the averaged loss, is the vector containing the neural network model parameters with size , and is the cardinality of the dataset . With (stochastic) gradient descent (SGD), the corresponding update rule to solve the problem above may be expressed as

where is the learning rate and is the global gradient vector.

In the case of federated learning, the global dataset is not available at the AP, i.e., STAs do not share their local dataset with the AP. In this scenario, to show why the gradient aggregation (Step 4 in FedSGD) solves the same optimization problem equivalently, we need to re-express the global gradient as a weighted summation of the local gradients :

Thus, by rewriting the update rule, we can show the following identity:

Thus, the update rule can be equivalently expressed as a *weighted average of the local gradients,* i.e., , leading to Step 4 of FedSGD. Notice that the weight of the th ED is based on the cardinality of the local dataset and the cardinality of the global dataset, i.e., . If all the STAs have the same number of data samples, the *weighted average* becomes an *arithmetic mean of the local gradients*.

In a typical neural network, there can be millions of model parameters, i.e., very large , leading to a large number of gradients (the number of gradients is equal to the number of learnable parameters). Hence, federated learning can cause significant traffic in a wireless network as millions of parameters need to be exchanged between the radios, especially when there are many STAs. This is known in the literature as the *communication bottleneck* of federated learning, which can be more pronounced in a bandlimited-communication network, such as WLAN. With OAC, the communication bottleneck can be addressed effectively if the local gradients, i.e., , can be aggregated over the channel.

  

1. (b)

Figure 2 (a) Communication followed by Computation; (b) Computation during Communication (OAC)

To realize the OAC scheme in realistic systems, one of the most critical challenges is to remove distortions to the transmitted signals from the fading in wireless channels, time-frequency synchronization errors, hardware impairments, e.g., carrier frequency offset (CFO), phase offset (PO), power amplifier non-linearity, and power imbalance. Unlike traditional data communication, in which the impact of most of those distortions can be mitigated at the receiver or by applying certain coding schemes to the information to be transmitted, the OAC would require pre-correcting the distortions to the signals before they are transmitted so that certain desired arithmetic operations can take place in the air. For instance, under a fading channel, the received symbol can be expressed as

where is the fading channel coefficient between th STA and AP, and is the noise, which shows that the received signal is not the desired sum, i.e., , under the fading channel, where may be a local gradient, i.e., an element of . To receive , the STAs need to precode their transmissions so that the parameters add up constructively in the complex plane (i.e., called *coherent* summation). Otherwise, the phase of the signals arriving at the AP receiver will not be the same, destroying the coherent summation. Thus, to achieve coherent summation, a primary challenge is to apply a precoder that counteracts the impact of fading channels, inevitable hardware impairments, and mismatch between the carrier frequencies of STAs and AP, including CFO and PO, on the coherent superposition. To date, none of any communication standard has been designed with the capability to precode the transmitted signal to enable OAC.

It is worth noting that coherent summation, which is the focus of this subclause, is one of the techniques to achieve OAC for FL. There have been several non-coherent techniques in the literature. One may refer to [8] for additional information.

#### KPIs

The metrics for assessing an OAC scheme differ from traditional communication metrics, such as the data rate, bit error rate, and block-error rate, since OAC aims to compute a function. Often, the performance of an OAC scheme can be assessed via an MSE analysis for an analog OAC, where mentioned above is a continuous variable. For a digital OAC scheme (where is discretized based on a constellation), the probability of computing a single function (or a set of functions) incorrectly can also be used as a metric since the image of the function consists of a set of discrete values. The computation rate, i.e., the number of functions calculated per real dimension, is another metric that can be used for evaluating the efficiency of an OAC scheme. One can also obtain application-specific metrics such as test accuracy and convergence rate for FL when it is used with an OAC scheme. Below are some examples:

* Data transmissions needed per training cycle
* Storage requirements
* Computational power
* Accuracy
* Communication overhead
* Learning quality

#### Requirements

To support the OAC for FL using coherent methods, the network needs to meet the following requirements

* Time and frequency synchronization
* Phase alignment
* Power balance in UL transmissions
* Sufficient tolerance to noise and interference
* Overhead trade-off

#### Technical Feasibility Analysis

##### Standards Impact

Similar to beamforming and sounding procedure, the standard needs to provide mechanisms in both PHY and MAC to enable and support OAC, thereby mitigate the distortion due to channels, CFO, PO, power imbalance, and other radio impairments. As an example, [9] introduced a procedure that would require transmission of NDPs in both UL and DL for phase aligned OAC.

##### Technical feasibility

The technical feasibility for OAC, assuming it is implemented to achieve coherent summation, in Wi-Fi networks mostly depends on the techniques to obtain the precoder mentioned in 2.2.1.1 with a certain accuracy. In this subclause, we demonstrate how each STA as an ED, with help from its associated AP, can precode its transmitted symbols where there are phase errors in both UL and DL. For other types of Tx/Rx mismatches and impairments, further study may be needed.

Consider a scenario with STAs and one AP, where all the STAs are triggered to transmit simultaneously. Let denote the parameter at the th STA. Suppose that the goal of the AP is to receive with signal superposition on the same time-frequency resources.



Figure 3 System model (narrow band)

By considering the channels between STAs and AP, CFO, and PO and a narrowband representation of the signals, the received signal at the AP can be expressed as

where is the received signal from the th STA at the AP, and it can be expressed as

where is the transmitted signal for aggregation, is the uplink channel, and and are the CFO and the PO with respect to the local oscillator of the AP with the carrier frequency of , respectively, as shown in Figure 3. According to the expression of , the phase of the received signal for th STA is a function of the uplink channel, CFO, and PO, and the phase rapidly changes over time as

Since the mismatch is STA specific, i.e., depending on the index , there is no way to correct the mismatches from the combined received signal .

To describe the procedure for correcting the PO, we denote the downlink channel between the AP and the th STA as for a narrowband representation. Thus, the received signal at the th STA can be expressed as

where is the transmitted signal from the AP. Hence, the phase of the received signal at the th STA is a function of the downlink channel, CFO, and PO, and, similar to , also rapidly changes as a function of time as

To mitigate PO, one method may be using phase-coded pilots in a sequence of DL and UL transmissions. Figure 4 depicts the basic idea. One may refer to [8] for additional information regarding the feasibility of phase-coded pilots.



Figure 4 Phase-coded pilot to achieve phase-coherent reception between AP and the th STA [9]

#### References

1. AbdulRahman *et al.,* Journey from Centralized to Distributed On‑Site Learning and Beyond, IEEE IoT J., 2020
2. Liu *et al.,* From Distributed ML to Federated Learning, arXiv, 2021
3. Yuan *et al.,* Decentralized FL: A Survey and Perspective, IEEE IoT J., 2024
4. TR 23.758 Study on Edge Computing
5. IEEE P2805.4 Standard for Edge Collaboration Protocols for Federated Learning
6. Yassir Nawaz, et al., Towards Efficient Communications in Federated Learning: A Contemporary Survey, arXiv, 2022
7. Nurul Huda Mahmood, *et al.,* Federated Learning for 6G: Paradigms, Taxonomy, Recent Advances and Insights, arXiv, Dec 2023
8. Alphan Şahin, et al., A Survey on Over‑the‑Air Computation, IEEE Communications Surveys & Tutorials, 2023
9. Alphan Şahin, et al., Feasibility Study of Phase Synchronization for Wireless Federated Learning over WLAN, 11-25/304r0

***Editor: The paragraphs above are proposed to be added to the AIML SC technical report.***

## Use cases that leverage AIML to enhance WLAN performance

###  Use case 2: AIML-enhanced PHY Operational Parameter Recommendations

#### Use case description

In a typical WLAN environment, STAs operate under various conditions influencing their energy consumption and network performance. Examples of such conditions include the STA’s capabilities, the current usage scenario (e.g., idle vs. active with background traffic vs. having active audio/video traffic), the coverage area, the motion state, the current network load, as well as regulatory constraints. Each of these factors plays a crucial role in determining the most effective PHY settings for energy efficiency and performance optimization.

Under the current IEEE 802.11 standard, APs can influence the STA’s PHY parameter selection, for instance setting upper and lower limits on the transmission power, which is done to ensure compliance with regulatory requirements as well as managing network interference.

Despite this, the ultimate decision on which parameters to employ for transmission lies with the STA, which makes this decision based on a set of criteria established by its specific implementation, allowing it to adapt to its unique needs and operational context.

The AIML PHY Operational Parameter Recommendations use case leverages the dynamic capabilities of AIML alllowing APs to assist STAs in dynamically and proactively selecting the optimal PHY parameters. These recommendations are based on the current conditions within the Basic Service Set (BSS) and Extended Service Set (ESS), such as network load and uplink quality, which the AP monitors as the receiving end of the STA transmissions. The information, not typically available to STAs, allows the AP to offer insights that can lead to more informed decision-making by the STA. When technologies like beamforming, MU-MIMO or OFDMA are being used, a STA PHY parameters change can impact also other STAs in the same BSS. In these scenarios, the AIML functionality can predict the impact of such changes and help the STA avoid disrupting changes and/or prepare other STA that may impacted to it.

Also, STA’s usage scenarios such as transitioning between sleep and active states dictate the selection of PHY parameters; in power-saving modes, STAs usually prioritize energy efficiency, whereas in performance-critical scenarios, maintaining high throughput and low latency is paramount. In such instances, predicting the impact of any changes to PHY parameters before their implementation helps minimizing the potential performance degradation.

On the other hand, the dynamic adjustment of PHY parameters plays a role in preserving the privacy of STAs, as unique sets of parameters can be used to fingerprint devices and therefore identify them unwittingly. One of the key privacy-enhancement strategies involves TX power randomization, which is designed to obfuscate the STA's identity and make it more challenging for unauthorized entities to track device activities through signal fingerprinting across EDP (Enhanced Data Privacy) epochs. This is particularly true in the context of the privacy-enhancement features discussed within the IEEE 802.11bi task group.

Given these complexities, there is a clear need for a mechanism where the AP proactively assists STAs in updating their PHY parameters. Such a mechanism enables the STAs to dynamically adjust their settings in real-time, accommodating the fluctuating conditions of the network environment. This proactive assistance ensures that STAs can maintain optimal performance and energy efficiency, aligning with both their immediate operational needs as well as infra-level performance optimizations.

#### KPIs

The KPIs considered in this use case are proposed as follows:

* Energy Efficiency:
	+ Reduction in power usage by STAs.
* Link Quality:
	+ Maintenance of acceptable link quality as defined by user-specified metrics such as retry rates and latency.
* Privacy:
	+ Reduction of the likeliness of STA identification across IEEE 802.11bi EDP epochs.

#### Requirements

The following potential features can be analyzed in this use case:

* Collection of reliable link statistics to inform recommendations.
* Capability of STAs to process and apply recommended PHY parameters.
* Integration of an AIML engine at the AP to generate parameter recommendations.

#### Technical Feasibility Analysis

The feasibility of this use case hinges on the availability of comprehensive link statistics and the capability of the AIML algorithms to process such data to generate accurate recommendations. The AP's ability to consider current BSS status and receiver-side statistics is crucial for recommendation accuracy.

##### Standard Impact

This use case requires extensions to existing spectrum management protocols, such as the addition of flags in dot11SpectrumManagement to advertise AP capabilities for PHY recommendations. It also necessitates defining new message formats for STA request and AP response interactions.

##### Technical feasibility

Depending on the specific implementation, the AIML capabilities can be either handled on each individual AP or managed by a central controller; the entity managing the model training would require sufficient computation capabilities to train and execute the implemented model.

The accuracy and longevity of recommendations are expected to improve with prolonged connectivity and stable link conditions.

#### Privacy Considerations

Simulations analyzing the effectiveness of frame anonymization techniques ([2] , [3] ) showed that Randomized and Changing MAC (RCM) alone is insufficient for protecting STA identities.

This is because various PHY and MAC features can still create a device’s fingerprint, enabling adversaries to de-randomize and track STAs across epoch boundaries.

To address this, the IEEE 802.11bi Task Group ([4] ) has discussed the concept of PHY parameter randomization, which aims to enhance privacy by altering the STA’s RF fingerprinting across Enhanced Data Privacy (EDP) epochs.

The "AIML PHY Operational Parameter Recommendations" use case provides a means to implement this privacy-enhancing feature. It also tackles concerns related to performance impacts, such as those arising from excessive power reduction, while ensuring optimal energy efficiency.

# Conclusions and Recommendations

## Conclusions

## Recommendations

# References

1. 11-22/987r0: AIML TIG Technical Report, November 2023
2. Domenico Ficara, Rosario G. Garroppo, and Jerome Henry. 2024. A Tutorial on Privacy, RCM and Its Implications in WLAN. IEEE Communications Surveys & Tutorials 26, 2 (2024), 1003--1040 <https://ieeexplore.ieee.org/document/10368019>
3. 11-24/1234r0: Analysis of Frame Anonymization Techniques
4. 11-24/550r1: Tx Power Randomization
5. 11-24/911r0: AIML PHY Operational Parameter Recommendations