Title

Blind Modulation Classification of Wi-Fi 6 and 5G signals for Spectrum Sensing

Authors

Kim, Byungjun
Mecklenbräuker, Christoph F
Gerstoft, Peter

Publication Date

2023-10-30

DOI

10.1145/3616388.3617527

Blind Modulation Classification of Wi-Fi 6 and 5G signals for Spectrum Sensing

Byungjun Kim*
University of California, San Diego
La Jolla, CA, USA
byk004@ucsd.edu

Christoph F. Mecklenbräuker
TU Wien
Vienna, Austria
cfm@tuwien.ac.at

Peter Gerstoft
University of California, San Diego
La Jolla, CA, USA
gerstoft@ucsd.edu

ABSTRACT

Classification of modulation of Wi-Fi 6 and 5G downlink (DL) user data signals for spectrum sensing is studied. First, the orthogonal frequency division multiplexing (OFDM) symbol duration and cyclic prefix (CP) length are estimated based on the cyclic autocorrelation function (CAF). We propose a feature extraction algorithm characterizing the modulation of OFDM signals based on the estimated parameters. The algorithm includes removing the effects of a synchronization error and converting the obtained feature into a 2D histogram of phase and amplitude. This histogram is input to a convolutional neural network (CNN)-based classifier. Our system works without knowledge of a carrier frequency, Wi-Fi preamble, or resource allocation of 5G physical channels. We evaluate the classifier's performance with data with various protocol-compliant configurations. Our classifier achieves at least 98% accuracy when SNR is above the value required for data transmission.

CCS CONCEPTS

• **Networks** → **Cognitive radios**; • **Computing methodologies** → *Neural networks.*

KEYWORDS

Modulation classification, OFDM, Wi-Fi, 5G, Spectrum sensing

ACM Reference Format:
Byungjun Kim, Christoph F. Mecklenbräuker, and Peter Gerstoft. 2023. Blind Modulation Classification of Wi-Fi 6 and 5G signals for Spectrum Sensing. In *Proceedings of the Int'l ACM Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM '23), October 30 – November 3, 2023, Montreal, Canada.*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3616388.3617527

1 INTRODUCTION

The growth of wireless communication technologies has necessitated efficient usage of the radio spectrum, a challenge that is being addressed with cognitive radio. A critical component of cognitive radio is intelligent spectrum sensing, which allows for a more precise characterization of spectrum usage and aids in better decision-making for spectrum allocation. Spectrum sensing encompasses signal detection [7], predicting future spectrum [27], and identifying modulation schemes. In this study, we focus on the classification of modulations of practical orthogonal frequency division multiplexing (OFDM) signals. This also enables applications such as channel quality estimation between a transmitter (TX) and a receiver (RX) for a spectrum sensor and wireless network troubleshooting.

OFDM has become essential in modern wireless communication systems, such as Wi-Fi 6 and 5G. In these systems, message bits are converted to digital symbols using modulation schemes such as quadrature phase shift keying (QPSK) and transmitted via data subcarriers. In OFDM, the multiple symbols are stacked in subcarriers within the frequency domain, so each OFDM time sample contains only a fraction of the information on the multiple frequency domain symbols. As a result, the modulation classifiers designed for single-carrier signals [19] cannot be applied directly to OFDM signals. Therefore, a precise modulation classification of Wi-Fi 6 and 5G signals requires additional processing beyond using raw time-domain samples as inputs.

A spectrum sensor must handle OFDM signals with diverse configurations without prior information, in contrast to user equipment connected to a wireless network. In Wi-Fi 6 and 5G systems, information about the user data transmission, including the modulation, is provided to the RX. However, since a spectrum sensor does not have prior knowledge of the type of signals it detects, it cannot deploy a protocol-specific procedure to obtain information about user data transmission. The parameters shaping OFDM signals, fast Fourier transform (FFT) size to generate inverse fast Fourier transform (IFFT) sequence, and cyclic prefix (CP) length, might be different even among OFDM signals with the same modulation scheme. Moreover, the carrier frequency configurations in 5G become increasingly diverse and data transmission might occupy only a part of channel bandwidth. As a result, estimation of these carrier frequency configurations is becoming increasingly difficult using transmission bandwidth and center frequency alone. Thus, a modulation classifier for spectrum sensing should estimate the modulation scheme using only the observed user data transmission without knowledge of carrier frequency.

We present a system to classify the modulation of the signals in Wi-Fi 6 [14] and 5G [2] for a spectrum sensing system. Without knowledge of the transmitter (TX) carrier frequency, Wi-Fi preamble, or 5G control information, our system deploys only the basic OFDM structure, IFFT sequence, and CP. The system includes the estimation of OFDM parameters: CP length and subcarrier spacing (SCS), which is directly related to the FFT size of the IFFT sequence. We focus on identifying modulation schemes used in the payload of Wi-Fi 6 signals and the physical downlink shared channel (PDSCH)

MSWiM '23, October 30-November 3, 2023, Montreal, QC, Canada

ACM ISBN 979-8-4007-0366-9/23/10.
https://doi.org/10.1145/3616388.3617527

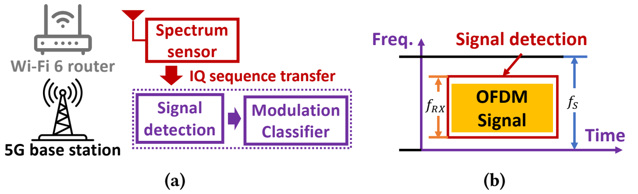
of 5G signals. Signals studied in this paper are single-input single-output (SISO). For 5G, they are in the frequency range 1 (FR1), whose frequency band is below 7.125 GHz.

For IFFT sequence and CP length estimation, the cyclic autocorrelation function (CAF) is deployed. The capability of CAF detecting repeated sequences as well as repetition periods enables the estimation of those parameters. We observe that symbol-level synchronization is not perfect if autocorrelation using CP is utilized only. Our preprocessing removes the effect of the synchronization error by using phase differences between phases of two adjacent OFDM symbols. The modulation classifier for Wi-Fi 6 and 5G signals should recognize high-order modulations such as 256 quadrature amplitude modulation (QAM) and 1024QAM since these state-of-the-art protocols include those schemes. We change the feature format to a histogram representing the distribution of the features so that the classifier can effectively capture high-order modulations characteristics.

Related work on modulation classification: Many papers address modulation classification for wireless communication signals [4, 9, 11, 12, 16, 19, 23, 24, 28]. The works in [4, 9, 11, 12, 24, 28] study modulation classification of OFDM signals and achieve at least 78% accuracy at 20 dB SNR for an AWGN channel. It is assumed that the inputs start from the first sample of OFDM symbol duration [4, 11, 12, 28], which is only possible by detecting Wi-Fi preamble or 5G synchronization signals properly. To apply this idea to a spectrum sensor, the sensor should follow protocol-specific procedures.

There are papers on OFDM modulation classification without the symbol-level synchronization assumption [9, 16, 23, 24] and the algorithms [9, 16, 24] are evaluated with hardware-generated data. However, their algorithms [9, 16, 24] are not evaluated with high-order modulations such as 256QAM or 1024QAM, as used in Wi-Fi 6. Moreover, since their classifier structures [9, 24] is designed to recognize a fixed set of modulations, the overall structure needs redesign to identify a new modulation scheme. The work [23] proposes the system to estimate SCS of OFDM signals and modulation of single-carrier signals jointly. Nonetheless, it does not estimate the modulation of OFDM signals. The neural network-based modulation classifier [19] studies how environmental change affects classification performance for only the single-carrier signals, not OFDM signals.

Related work on sniffing OFDM signals: For spectrum sensing, modulation might be identified by sniffing control data used to notify RX. The work [5, 6, 10, 17] tried to overhear Long Term Evolution (LTE) signals. LTEye [17] and OWL [5] decodes PHY DL control channel (PDCCH) data for LTE network monitoring. LTESniffer [10] decode sniffed both user and control data using PDCCH decoder FALCON [6]. FALCON overcomes the limitation of LTEye and OWL, which require more than 97% decoding accuracy. In LTE, a starting symbol of PDCCH in a slot is always the first symbol in a slot and it is different from 5G, where the PDCCH starting symbol in a slot can be any symbol in a slot and its information is notified with radio resource control (RRC) signals. Accordingly, it is not straightforward to generalize LTE PDCCH sniffer to 5G. Eavesdropping PDCCH data of 5G signals [21] can deal with the signal with diverse 5G configurations, but is vulnerable to configuration changes since it takes a few minutes to learn a new PDCCH configuration. The authors of [18] study sniffing Wi-Fi probe request packets, which is for mobile devices to broadcast the existence of themselves. They build a hardware model for a sniffer and test with real Wi-Fi probe request packets. However, the probe request packets are simpler than those for user data communication thus not straightforward to deploy this system for our target signal.



(a)

(b)

Figure 1: (a) System model to capture DL Wi-Fi 6 and 5G signals and (b) Signal detection scenario.

To summarize, the main contributions of the paper are:

- **OFDM parameter estimation for up-to-date protocols:** We have applied the OFDM parameter estimation method with CAF [25] to Wi-Fi 6 and 5G signals to estimate SCS and CP length.
- **Feature extraction without symbol-level synchronization:** Only with estimated values of SCS and CP length, our system builds the features characterizing modulation of OFDM signals. The proposed feature extraction algorithm is designed to be resilient to symbol-level synchronization errors caused by using CP only.
- **Modulation classification without control information:** For spectrum sensing, control information might not be accessible. We show that the proposed classification system robustly works with diverse configurations with the evaluation of hardware-generated data without knowledge of the information.

2 SYSTEM OBJECTIVE

We aim to build a modulation classifier using in-phase and quadrature (IQ) samples of SISO Wi-Fi 6 and FR1 5G DL signal for spectrum sensing. The system scenario is described in Fig. 1a. There is a Wi-Fi 6 or 5G TX transmitting its signal to an RX. The spectrum sensor continuously senses the spectrum by generating IQ samples with sampling rate f_S and transfers those samples to a signal detection algorithm. Using IQ samples captured by a receiver antenna, the signal detection algorithm detects the duration and frequency band where the OFDM signal is located and extracts IQ samples corresponding to the detected OFDM signal, described as the blue rectangle in Fig. 1b. We assume the accurate signal detection of Wi-Fi 6 or 5G signals and a single modulation scheme is used for data communication in one detected OFDM signal.

The IQ samples from the spectrum sensor sampled with rate f_S are resampled to f_{RX}, 20 MHz. We only consider Wi-Fi 6 signals with 20 MHz channel bandwidth and 5G signals with a PDSCH bandwidth from 15 to 20 MHz. Thus, 20 MHz sampling rate can let

Table 1: Variable definitions

Variable	Definition (unit)
f_{TX}	TX sampling rate (Hz)
f_{RX}	Sampling rate of a system input sequence (Hz)
Δf_{SCS}	Subcarrier spacing (Hz)
T_{IFFT}	IFFT sequence duration (s)
N_{FFT}	FFT size used to generate IFFT sequence
T_{CP}	CP duration (s)
N_{CP}	Number of time samples in CP for one OFDM symbol
$y[n]$	Received time-domain sequence after resampling to 20 MHz
$y'[n]$	5G time-domain sequence after resampling to 30.72 MHz
$y^i[n]$	Received time-domain IFFT sequence for the ith OFDM symbol
$Y^i[k]$	Received symbol in subcarrier k for the ith OFDM symbol
$(\mathcal{S}\times\mathcal{S})$	Number of bins of a classifier input

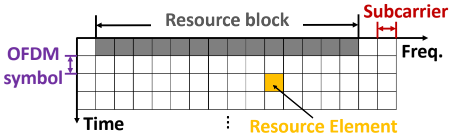


Figure 2: Example 5G resource grid.

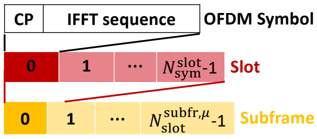

CP
IFFT sequence
OFDM Symbol
0
1
⋯
$N_{\mathrm{sym}}^{\mathrm{slot}}$-1
Slot
0
1
⋯
$N_{\mathrm{slot}}^{\mathrm{subfr},\mu}$-1
Subframe

Figure 3: 5G subframe structure.

the resampled IQ sequence encompass the OFDM signal in our scenario. Extending the analysis to different transmission bandwidth ranges is straightforward. These resampled IQ samples, denoted by $y[n]$, are taken as inputs of the feature extraction algorithm, as elaborated in Sec. 3 in detail.

2.1 Wi-Fi 6 PHY layer

Wi-Fi 6 supports the high-efficiency (HE) transmission format as well as earlier formats, such as non-high throughput (non-HT), high throughput (HT), and very high throughput (VHT) formats. Table 2 summarizes the parameters that configure the payload of the Wi-Fi frame for each Wi-Fi format. In HE format, the number of subcarriers is increased because the subcarrier spacing (SCS, denoted as Δf_{SCS}) is one-fourth of that of the previous transmission formats. Over time, the Wi-Fi standard has evolved and several options for the CP duration are available.

2.2 5G DL PHY layer

The 5G downlink (DL) resource structure and its associated terminology is illustrated in Fig. 2. A resource element (RE) represents the smallest unit which carries data, encompassing a single OFDM symbol in the time domain and a single subcarrier in the frequency domain. A resource block (RB) is the smallest radio resource that can be allocated and refers to one OFDM symbol in the time domain and 12 subcarriers in the frequency domain.

Figure 3 shows the 5G subframe structure in the time domain. An OFDM symbol in 5G is comprised of both a CP and an inverse fast Fourier transform (IFFT) sequence. The number of symbols within a single slot ($N_{\mathrm{sym}}^{\mathrm{slot}}$) varies in accordance with the CP length. There are a normal and an extended CP option in the transmission format. When a normal CP is used then $N_{\mathrm{sym}}^{\mathrm{slot}} = 14$, otherwise $N_{\mathrm{sym}}^{\mathrm{slot}} = 12$. The SCS, the distance between two adjacent subcarries in OFDM systems, denoted by μ, determines the number of slots within a single subframe, $N_{\mathrm{slot}}^{\mathrm{subfr},\mu}$. There are five SCS options in 5G, but we consider only three cases, namely 15, 30, 60 kHz, which are available in FR1. These SCS values correspond to $\mu = 0, 1, 2$, respectively, and the number of slots in a subframe for each SCS is computed as $N_{\mathrm{slot}}^{\mathrm{subfr},\mu} = 2^{\mu}$.

The structural parameters which define the 5G frame are listed in Table 3. The length of an IFFT sequence, T_{IFFT}, is:

$$T_{\mathrm{IFFT}} = N_{\mathrm{FFT}}/f_{\mathrm{TX}} = 1/\Delta f_{\mathrm{SCS}}. \tag{1}$$

There is a one-to-one correspondence between T_{IFFT} and Δf_{SCS} (1). Under the normal CP option, CP is longer than that in other symbols, every 0.5 ms, or equivalently, $7\cdot 2^{\mu}$ OFDM symbols in OFDM symbol unit, called long CP. There is no long CP in the extended CP option, so T_{CP} is uniform. The transmission rate of 5G signals is a power of 2 times 15 kHz and 30.72 MHz is an example of 5G transmission rate. N_{FFT} and N_{CP} values are arranged when f_{TX} is 30.72 MHz, the value used in our evaluation.

In addition to PDSCH, there exist other physical (PHY) channels that but serve specific functions although not carrying user data. For instance, PDCCH conveys downlink control information (DCI), which contains information required to decode PDSCH data such as modulation and coding scheme (MCS). Each of these channels utilizes predefined single-type modulation, see Table 4.

Compared to Wi-Fi, which has a predefined configuration of data, pilot, and null subcarriers, 5G resource configuration for PHY channels is flexible. Instead, the 5G system has a network dedicated to exchanging information on how data packets are forwarded, called the control plane, in addition to the network for data transmission, called the user plane. An example of data transferred over the control plane is RRC signals. Information on the starting OFDM symbol of PDCCH and channel state information-reference signal (CSI-RS) is notified to an RX with RRC signals via control plane [2].

3 PROPOSED ALGORITHM

High-level procedures to build features characterizing the modulations of Wi-Fi 6 and 5G signals are illustrated in Fig. 4 and explained in Sec. 3.1 and 3.2. The output of this flowchart is taken as an input to the neural network model, described in Sec. 3.3.

Table 2: Parameters for different formats of Wi-Fi

	Non-HT format	**HT format**	**VHT format**	**HE format**
Channel bandwidth	20 MHz	{20, 40} MHz	{20, 40, 80, 160} MHz	{20, 40, 80, 160} MHz
T_{IFFT}	$3.2\,\mu$s	$3.2\,\mu$s	$3.2\,\mu$s	$12.8\,\mu$s
T_{CP}	$0.8\,\mu$s	$\{0.4, 0.8\}\,\mu$s	$\{0.4, 0.8\}\,\mu$s	$\{0.8, 1.6, 3.2\}\,\mu$ss
Modulations	BPSK, QPSK, 16QAM, 64QAM	BPSK, QPSK, 16QAM, 64QAM	BPSK, QPSK, 16QAM, 64QAM, 256QAM	BPSK, QPSK, 16QAM, 64QAM, 256QAM, 1024QAM

Table 3: 5G frame structure parameters

{SCS (kHz), CP option}	**{60, Normal}**	**{60, Extended}**	**{30, Normal}**	**{15, Normal}**
T_{IFFT}	$16.17\,\mu$s	$16.67\,\mu$s	$33.33\,\mu$s	$66.67\,\mu$s
{Short, long} T_{CP}	$\{1.17, 1.69\}\,\mu$s	$\{4.17, -\}\,\mu$s	$\{2.34, 2.86\}\,\mu$s	$\{4.69, 5.21\}\,\mu$s
Long CP period (in symbol)	28	-	14	7
N_{FFT} **when** f_{TX} **= 30.72 MHz**	512	512	1024	2048
{Short, long} N_{CP} **when** f_{TX} **= 30.72 MHz**	{36, 52}	128	{72, 88}	{144, 160}

Table 4: Modulations used for 5G physical channels

Physical channel	**PDSCH**	**PSS/SSS**	**PDCCH**	**CSI-RS**	**PBCH**	**PDSCH-PTRS**	**PDSCH-PTRS**
Modulation	QPSK, 16QAM, 64QAM, 256QAM	BPSK	QPSK	QPSK	QPSK	QPSK	QPSK

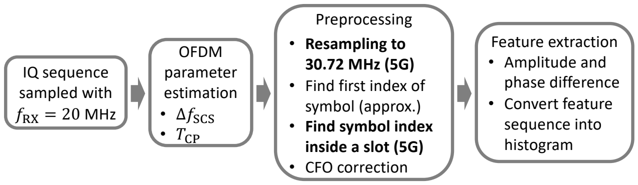

IQ sequence sampled with f_{RX} = 20 MHz
OFDM parameter estimation
• Δf_{SCS}
• T_{CP}
Preprocessing
• Resampling to 30.72 MHz (5G)
• Find first index of symbol (approx.)
• Find symbol index inside a slot (5G)
• CFO correction
Feature extraction
• Amplitude and phase difference
• Convert feature sequence into histogram

Figure 4: Flow chart of proposed feature extraction algorithm.

3.1 OFDM parameter estimation

Prior to building the features which characterize modulation, it is necessary to estimate two essential OFDM parameters of OFDM signals, SCS and CP length. To estimate these parameters, we use CAF, a Fourier-series cofficient of autocorrelation function.

$$\mathcal{R}_{yy}(\alpha,\tau) = \sum_{n=-\infty}^{\infty} \mathcal{R}_{yy}(n,\tau)\mathrm{e}^{-j2\pi\alpha n}. \tag{2}$$

CAF is used to extract a repeated pattern presented in wireless signals [8, 13, 25]. A variant of the CAF estimator presented in [25] is deployed here,

$$\hat{\mathcal{R}}_{yy}(\alpha,\ell) = \frac{1}{\mathcal{N}}\sum_{n=0}^{\mathcal{N}-1}\left\{\sum_{i=0}^{l-1} y(n+i)y^*(n+i+\ell)\right\}\mathrm{e}^{-j2\pi\alpha n}. \tag{3}$$

One sample of our estimator is computed as the autocorrelation with delay ℓ. It differs from the estimator in [25], where only two samples are used to compute one estimator sample. This modification aims to make peaks more distinct. We set $l = 8$ corresponding to the shortest CP length.

CP in OFDM symbols causes a sequence to be repeated at both ends of each symbol. The distance between starting indices of the two repeated sequences located at both ends of an OFDM symbol is T_{IFFT} or $N_{\mathrm{FFT}}(f_{\mathrm{RX}}/f_{\mathrm{TX}}) = f_{\mathrm{RX}}/\Delta f_{\mathrm{SCS}}$, depending on whether it is in time units or time sample units, respectively. This repetition makes the CAF estimator at $\alpha = 0$ have a peak at $\ell = f_{\mathrm{RX}}/\Delta f_{\mathrm{SCS}}$. T_{CP} is also estimated with the CAF estimator, $\hat{\mathcal{R}}_{yy}(\alpha, f_{\mathrm{RX}}/\Delta f_{\mathrm{SCS}})$. Since $\sum_{i=0}^{l-1} y(n+i)y^*(n+i+\tau)$ in Eq. (3) has peaks with the period of $f_{\mathrm{RX}} \cdot (T_{\mathrm{CP}} + 1/\Delta f_{\mathrm{SCS}})$, it is expected of $\hat{\mathcal{R}}_{yy}\,(\alpha, f_{\mathrm{RX}}/\Delta f_{\mathrm{SCS}})$ to have a large amplitude at $\alpha = 1/\{f_{\mathrm{RX}} \cdot (T_{\mathrm{CP}} + 1/\Delta f_{\mathrm{SCS}})\}$.

In our scenario, there are five candidates for ℓ values, $\boldsymbol{\ell}_C = \{64, 256, 333, 667, 1333\}$, each of which corresponds to an IFFT sequence length for a given SCS at $f_{\mathrm{RX}} = 20$ MHz. IFFT sequence length is estimated as:

$$T_{\mathrm{IFFT}} = \ell'/f_{\mathrm{RX}} \quad \text{s.t.} \quad \ell' = \underset{\ell\in\boldsymbol{\ell}_C}{\arg\max}\left|\hat{\mathcal{R}}_{yy}(0,\ell)\right| \tag{4}$$

When the estimated T_{IFFT} corresponds to that of Wi-Fi 6 or 60 kHz SCS NR, where multiple CP options are available, CP length is further estimated as:

$$T_{\mathrm{CP}} = \frac{1}{f_{\mathrm{RX}}}\left(\frac{1}{\alpha'} - \ell'\right) \quad \text{s.t.} \quad \alpha' = \underset{\alpha\in\boldsymbol{\alpha}_C^{\ell'}}{\arg\max}\left|\hat{\mathcal{R}}_{yy}(\alpha,\ell')\right| \tag{5}$$

where $\boldsymbol{\alpha}_C^{\ell'}$ denotes a set of possible values of $\alpha = 1/\{\ell' + (f_{\mathrm{RX}} \cdot T_{\mathrm{CP}})\}$, given ℓ'.

3.2 Feature extraction

The motivation behind our proposed feature extraction lies in the observation that when a sampled time-domain sequence is contained within a single OFDM symbol, the Fast Fourier Transform

(FFT) of that sequence yields the original symbols with a phase drift that scales linearly with subcarrier index k and synchronization error Δn, as shown in

$$\begin{aligned} Y^i_{\Delta n}[k] &\triangleq \mathcal{F}\left(y^i[n-\Delta n]\right) = \sum_{n=0}^{N_{\text{FFT}}-1} y^i[n-\Delta n]\mathrm{e}^{-j2\pi nk/N_{\text{FFT}}} \\ &= Y^i[k]\mathrm{e}^{-j2\pi\Delta nk/N_{\text{FFT}}}. \end{aligned} \tag{6}$$

In order to build a feature characterizing modulation based on this property, two objectives must be achieved: first, sampling a sequence that is fully contained within an OFDM symbol, and second, removing the phase drift caused by synchronization errors.

Utilizing the knowledge of N_{CP} and N_{FFT}, the CP position is determined through autocorrelation analysis,

$$R_{yy}(n, N_{\text{FFT}}) = \frac{1}{N_{\text{CP}}} \sum_{i=0}^{N_{\text{CP}}-1} y[n+i]y^*[n+i+N_{\text{FFT}}]. \tag{7}$$

The position of CP is indicated by the peaks in $|R_{yy}(n, N)|$. To locate a peak, we search for a sample whose amplitude is larger than both of its neighboring samples while ensuring that the minimum distance between two adjacent peaks is 90% of the OFDM symbol duration (i.e., 288-time sample indices for HE format with 3.2 μs CP), to avoid selecting undesired local peaks. We compute their remainders divided by the sample number of the OFDM symbol duration ($N_{\text{CP}} + N_{\text{FFT}}$) over multiple OFDM symbols. The median of those remainders is determined as the first index of the OFDM symbol, denoted as p. Noise and varying amplitudes of time samples can introduce small errors in the estimated CP position. To reliably sample the sequences contained in a single OFDM symbol, we deploy the sequence $\{y[p+N_{\text{CP}}/2], y[p+N_{\text{CP}}/2+1], \cdots, y[p+N_{\text{CP}}/2+N-1]\}$. This sequence is entirely within a single OFDM symbol unless the estimation error of p is larger than $N_{\text{CP}}/2$.

We demonstrate that $Y^i_{\Delta n}[k]$ exhibits a phase drift modeled by $\mathrm{e}^{-j2\pi\Delta nk/N}$, while maintaining its amplitude $Y^i[k]$ (6). We remove this phase drift due to synchronization errors by computing the phase differences between successive symbols in the same subcarrier k as:

$$\begin{aligned} \Delta\angle Y^i_{\Delta n}[k] &\triangleq \angle Y^{i+1}_{\Delta n}[k] - \angle Y^i_{\Delta n}[k] \\ &= \angle\left\{Y^{i+1}[k]\mathrm{e}^{-j2\pi\Delta nk/N}\right\} - \angle\left\{Y^i[k]\mathrm{e}^{-j2\pi\Delta nk/N}\right\} \\ &= \angle Y^{i+1}[k] - \angle Y^i[k]. \end{aligned} \tag{8}$$

Despite the lack of knowledge about Δn, sequences with constant Δn can be obtained by adjusting the interval between the starting indices of two sampled sequences to be one OFDM symbol. The feature used to identify the modulation type is $Y^i_f[k] \triangleq |Y^i_{\Delta n}[k]|\mathrm{e}^{j\Delta\angle Y^i_{\Delta n}[k]}$. The null subcarrier symbols is eliminated by discarding symbols with the N_{null} smallest amplitudes.

In protocol-compliant reception, the Wi-Fi preamble and 5G PDSCH-phase tracking reference signal (PDSCH-DMRS) are deployed for CFO estimation. However, since they are not accessible to a spectrum sensor, the CP in each OFDM symbol is used for CFO estimation, i.e.,

$$\angle\left(\frac{y(p+N_{\text{FFT}}+i)}{y(p+i)}\right) = 2\pi\Delta f_c/\Delta f_{\text{SCS}}, \tag{9}$$

where $y(p+i)$ is in CP. We use $i \in \{\lfloor N_{\text{CP}}/4\rfloor, \cdots, \lceil 3N_{\text{CP}}/4\rceil\}$ so that the sequence $y(p+i)$ are entirely within CP unless estimation error of p is larger than $N_{\text{CP}}/4$. If the absolute value of the CFO is larger than $\Delta f_{\text{SCS}}/2$, CFO cannot be accurately estimated due to aliasing. It is discussed in Sec. 3.2.1 in detail.

3.2.1 Additional procedures for 5G signal. To build a modulation feature for 5G, 5G characteristics distinct from those of Wi-Fi, including a different transmission rate, long CP, and flexible usage of subcarriers, should be considered. First, the transmission rate of 5G signals is not $f_{\text{RX}} = 20$ MHz, but is the form of a power of 2 times 15 kHz. Hence, if the signal is classified as 5G, we resample the sequence to 30.72 MHz = $2048 \cdot 15$ kHz, the smallest sampling frequency above 20 MHz. N_{FFT} and N_{CP} with 30.72 MHz sampling rate for each Δf_{SCS} are arranged in the last two rows in Table 3.

In the case of the normal CP option, there is a long CP every 0.5 ms, which is slightly longer than that of other OFDM symbols. Long CP breaks the assumption of the uniform OFDM symbol durations, which is required by the method to find the first indices of OFDM symbols and to build $Y^i_f[k]$. Specifically in building $Y^i_f[k]$, maintaining the fixed interval does not guarantee the constant Δn over multiple OFDM symbols. Therefore, long CP also should be located when finding the first index of the OFDM symbol.

```
Algorithm 1: Finding first index of long CP
   Data: ($y'[n]$ of length (3 ms + 3 OFDM symbols)), $\mu$
   Result: firstIndexLongCP= $q + \text{symLongCP} \cdot (N_{\text{FFT}} + N_{\text{CP}})$
1  $m = 7 \cdot 2^{\mu}$, $N_{\text{FFT}} = 512 \cdot 2^{2-\mu}$, $N_{\text{CP}} = 18 \cdot 2^{2-\mu}$, $i = 0$;
2  while $i \leq 5$ do
3      $y'_i[n] = \{y'[(30.72 \cdot 10^6) \cdot (0.5 \cdot 10^{-3}) \cdot i], \cdots, y'[(30.72 \cdot 10^6) \cdot (0.5 \cdot 10^{-3}) \cdot (i+1) + 3 \cdot (N_{\text{FFT}} + N_{\text{CP}})]\}$;
4      Find peaks $\{p'_{i0}, \cdots, p'_{i(m+1)}\}$ with $y'_i[n]$ using
         autocorrelation $|R_{y'_i y'_i}(n, N_{\text{FFT}})|$ and peak locating
         function explained in Sec. 3.2;
5      $p_{ij} = \text{mod}(p'_{ij}, N_{\text{FFT}} + N_{\text{CP}})$, $i = i + 1$;
6  end
7  $\Delta p_j = \text{Mean}(\{p_{0(j+1)} - p_{0(j-1)}, \cdots, p_{5(j+1)} - p_{5(j-1)}\})$;
8  $\{\Delta p_{k_0}, \cdots, \Delta p_{k_{m-1}}\} = \text{sortDescending}(\{\Delta p_j\})$;
9  symLongCP = $\arg\max_{k_q} \text{Var}(\{p_{0k_q}, \cdots, p_{5k_q}\})$ where
     $q \in \{0, 1\}$;
10 $q_{ij} = \begin{cases} p_{ij} & \text{if } j \leq \text{symLongCP} \\ p_{ij} - 16 & \text{otherwise} \end{cases}$
     $q_j = \text{Median}(\text{Mean}(\{q_{0j}, \cdots q_{5j}\}))$ where
     $j \in \{0, 1, \cdots, m-1\} - \{\text{symLongCP}\}$;
```

Algorithm 1 explains the detailed steps to estimate the first index of OFDM symbol with long CP. $y'_i[n]$ in line 3 is a sequence cropped to be as long as (0.5 ms + 3 OFDM symbols). In line 4, we find $m+2$ peaks from $y'_i[n]$ using autocorrelation $|R_{y'_i y'_i}(n, N_{\text{FFT}})|$, where m denotes the number of OFDM symbols in 0.5 ms given μ. The difference between the remainders of two peaks separated by two OFDM symbols divided by OFDM symbol duration, Δp_j, is computed as the average of $p_{i(j+1)} - p_{i(j-1)}$ over i. We expect that Δp_j is the largest when p_j corresponds to long CP. For a more reliable estimation of a long CP, we add one additional criterion. In line 8, we choose the two candidates k_0 and k_1 that give Δp_{k_i} the two largest values. We select k_q where the samples $\{p_{0k_q}, \cdots p_{5k_q}\}$ has the larger variance between two candidates of k_q. This is because we expect that $\{p_{0j}, \cdots, p_{5j}\}$ has the largest variance if p_{ij} corresponds to long CP since long CP makes $|R_{y'_i y'_i}(n, N_{\text{FFT}})|$ a plateau with certain width, not one sharp peak caused by non-long CP. Using estimated **firstIndexLongCP**, we put an additional 16 samples delay at the OFDM symbol with long CP while extracting the feature $Y_f^i[k]$ to maintain uniform Δn. The number of 16 samples comes from the difference between long CP and non-long CP with 30.72 MHz sampling rate.

In contrast to Wi-Fi 6 signals, some subcarriers might not be used for transmission in the midst of transmission. If no transmission is made in $Y^i[k]$ or $Y^{i+1}[k]$, their phases are random, and $\Delta\angle Y_{\Delta n}^i[k]$ cannot be represented as the phase difference. Therefore, we set the threshold for the amplitude, denoted as β, to check whether the PE is being used for transmission. Only when the amplitudes of both subcarrier symbols in adjacent OFDM symbols are higher than the threshold, this feature is used.

The discrepancy between the center frequency of TX and that of received IQ samples of 5G signals might be much larger than for Wi-Fi. This is because payload in Wi-Fi covers the entire channel bandwidth unless OFDMA is used. In contrast, PDSCH in 5G might use only the part of channel bandwidth so the center frequency of PDSCH might be different from that used for transmission. Thus, the discrepancy is solely from hardware imperfection in Wi-Fi. For a Wi-Fi link operating at $f_c = 5\,\text{GHz}$ and a frequency tolerance of 1 ppm for commercial-off-the-shelf temperature-compensated crystal oscillators [20] on both sides of the Wi-Fi link, the worst-case CFO is $\Delta f_c = 2f_c \cdot 10^{-6} = 10\,\text{kHz}$. However, in 5G, CFO can escalate to an MHz scale if we consider the center frequency of transmission bandwidth to be carrier frequency. If the method presented earlier in this subsection is employed, the difference could result in an inaccurate estimation of CFO due to aliasing. Even in absence of noise, it is only possible to measure Δf_c accurately up to $\Delta f_{\text{SCS}}/2$, since $\Delta f_c + j \cdot \Delta f_{\text{SCS}}$ cannot be distinguished from each other, where $j \in \mathbb{Z}$. The provided algorithm makes the corrected CFO a multiple of Δf_{SCS}, not a zero.

However, the CFO correction algorithm is still deployed for feature extraction. This is because even though this method cannot find the exact CFO, it can recover the orthogonality among subcarriers. The CFO effect in our feature can be represented as:

$$\begin{aligned}
Y_{\Delta n}^i[k] &= \sum_{n=0}^{N_{\text{FFT}}-1} y[n-\Delta n]\mathrm{e}^{-j2\pi n(\Delta f_c/f_{\text{TX}}+k/N_{\text{FFT}})} \\
&= Y^i\left[k + N_{\text{FFT}}\Delta f_c/f_{\text{TX}}\right]\mathrm{e}^{-j2\pi\Delta n(k/N_{\text{FFT}}+\Delta f_c/f_{\text{TX}})} \\
Y_{\Delta n}^{i+1}[k] &= Y^{i+1}\left[k + N_{\text{FFT}}\Delta f_c/f_{\text{TX}}\right] \times \\
&\quad \mathrm{e}^{-j2\pi(\Delta nk/N_{\text{FFT}}+(\Delta n+(N_{\text{FFT}}+N_{\text{CP}})\Delta f_c/f_{\text{TX}})} \\
\Rightarrow \Delta\angle Y_{\Delta n}^i[k] &= \angle Y^{i+1}\left[k + \Delta f_c/\Delta f_{\text{SCS}}\right] - \angle Y^i\left[k + \Delta f_c/\Delta f_{\text{SCS}}\right] \\
&\quad - 2\pi\Delta f_c(1/\Delta f_{\text{SCS}} + T_{\text{CP}}).
\end{aligned} \tag{10}$$

To maintain orthogonality of $\angle Y_{\Delta n}^i[k]$ across k, $\Delta f_c/\Delta f_{\text{SCS}}$ should be an integer. We have demonstrated that after the CFO correction using CP, the CFO value can be expressed as $j \cdot \Delta f_{\text{SCS}}$, which renders $\Delta f_c/\Delta f_{\text{SCS}}$ to be an integer. Consequently, the phase of our feature becomes the sum of a phase difference of originally transmitted symbols and a phase caused by CFO. Since $\Delta\angle Y_{\Delta n}^i[k]$ in (10) contains T_{CP} term, the CFO effect on $\Delta\angle Y_{\Delta n}^i[k]$ is different when OFDM symbol $i-1$ is an OFDM symbol with long CP. To make the CFO effect uniform in the feature, $\Delta\angle Y_{\Delta n}^i[k]$ where OFDM symbol $i-1$ is not used for building the feature.

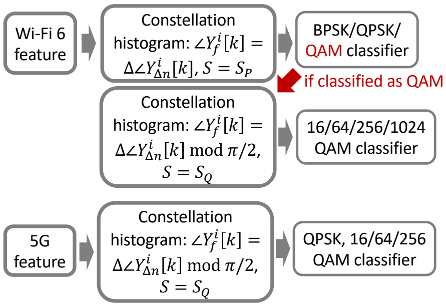

Wi-Fi 6 feature
Constellation histogram: $\angle Y_f^i[k] = \Delta\angle Y_{\Delta n}^i[k]$, $S = S_P$
BPSK/QPSK/ QAM classifier
if classified as QAM
Constellation histogram: $\angle Y_f^i[k] = \Delta\angle Y_{\Delta n}^i[k] \bmod \pi/2$, $S = S_Q$
16/64/256/1024 QAM classifier
5G feature
Constellation histogram: $\angle Y_f^i[k] = \Delta\angle Y_{\Delta n}^i[k] \bmod \pi/2$, $S = S_Q$
QPSK, 16/64/256 QAM classifier

Figure 5: Flow chart of proposed classifier system.

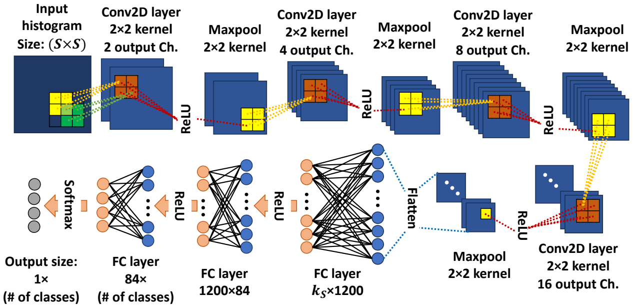

Input histogram Size: $(S\times S)$
Conv2D layer 2×2 kernel 2 output Ch.
ReLU
Maxpool 2×2 kernel
Conv2D layer 2×2 kernel 4 output Ch.
ReLU
Maxpool 2×2 kernel
Conv2D layer 2×2 kernel 8 output Ch.
ReLU
Maxpool 2×2 kernel
Conv2D layer 2×2 kernel 16 output Ch.
ReLU
Maxpool 2×2 kernel
Flatten
FC layer k_S×1200
ReLU
FC layer 1200×84
ReLU
FC layer 84× (# of classes)
Softmax
Output size: 1× (# of classes)

Figure 6: CNN-based modulation classifier structure.

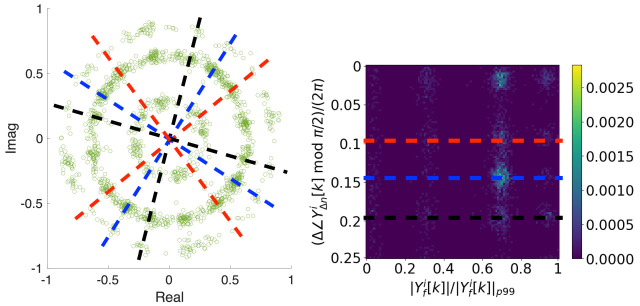
Table 5: DL model parameters

Batch size	32
Learning rate	$5 \cdot 10^{-5}$
Epochs	200
Loss	Cross-entropy

The features may contain the effect of other PHY channels which use modulations other than those used by PDSCH. It is impossible to perfectly filter out the effect because information about which REs were used for which PHY channels is not accessible. However, since the modulations of other PHY channels are either BPSK or QPSK, the constellation diagram of the features is only affected by PDSCH modulation. Thus, the distribution of phase differences is still an intrinsic characteristic of PDSCH modulation.

3.3 Input of neural network classifier

The obtained feature $Y_f^i[k]$ goes through two preprocessing steps to become input to the classifier. 1) instead of $\Delta\angle Y_{\Delta n}^i[k]$, $\Delta\angle Y_{\Delta n}^i[k]$ modulo $\pi/2$ is used as a phase of $Y_f^i[k]$. A constellation diagram of



(a)

(b)

Figure 7: Measured 16QAM features at SNR= 25 dB with 5G data: (a) Scatterplot of $Y^i_f[k]$ and (b) Corresponding histogram of $|Y^i_f[k]|/|Y^i_f[k]|_{\text{p99}}$ and $(\Delta\angle Y^i_{\Delta n}[k]$ mod $\pi/2)/2\pi$.

every target modulation and corresponding features $Y^i_f[k]$ without noise are symmetric with $\pi/2$. Thus, $\Delta\angle Y^i_{\Delta n}[k]$ modulo $\pi/2$ is used as a phase of our feature to characterize a modulation. For Wi-Fi 6 signals, BPSK cannot be distinguished from QPSK if $\Delta\angle Y^i_{\Delta n}[k]$ modulo $\pi/2$ is used. Thus, an additional classifier with the original phase as an input is used to distinguish BPSK and QPSK from the high-order QAM modulations, see Fig. 5. 2) A 2D histogram of the normalized amplitude of the features $|Y^i_f[k]|/|Y^i_f[k]|_{\text{p99}}$, where $|Y^i_f[k]|_{\text{p99}}$ denotes 99% percentile of $|Y^i_f[k]|$ in a single data, and the phases $\angle Y^i_f[k]/2\pi$, as an input for the classifier. Each bin in the histogram is normalized by dividing it by the number of features in a single data. To remove outliers, $Y^i_f[k]$ whose amplitude is larger than $|Y^i_f[k]|_{\text{p99}}$ was not included in the histogram.

The overall structure and the parameter of the classifier with the histogram as input are summarized in Fig. 5 and Table 5. The neural network structure used for each classifier is described in Fig. 6. To identify BPSK and QPSK, $\mathcal{S} = \mathcal{S}_P$, and the third Maxpool layer is not used.

Figure 7a shows a scatterplot of the IQ data of $Y^i_f[k]$ and Fig. 7b the corresponding 2D histogram for 5G 16QAM data with $\Delta\angle Y^i_{\Delta n}[k]$ modulo $\pi/2$. $\angle Y^i_f[k]$ on the red and black dashed lines are the noise-free phase differences between two 16QAM symbols. Blue dashed lines are from the phase differences between BPSK or QPSK symbols of the PHY channel other than PDSCH. The noise-free phase difference values are (odd integer)$\cdot\pi/4$ and shifted further by CFO. The red, blue, and black dashed lines in Fig. 7a correspond to the red, blue, and black dashed lines in Fig. 7b, respectively. Fig. 7a and Fig. 7b show that symbols are densely located at the points in the dashed lines, which is consistent with our expectations.

An advantage of using a histogram is that they are invariant to the length of $Y^i_f[k]$. This enables a neural network with a fixed structure to handle signals of any duration. This property is useful when dealing with 5G features where the number of samples of $Y^i_f[k]$ is unknown due to unused resources. Moreover, in a histogram input, the effect of CFO estimation error caused by aliasing (10) is a movement along y-axis of the histogram as far as orthogonality of $\angle Y^i_{\Delta n}[k]$ across k holds. The neural network can be trained to identify histogram movements along y-axis as a single class.

Table 6: Data generation parameters

SNR	[5, 40] dB in steps of 5 dB
Carrier frequency	2.4 GHz (Wi-Fi 6), 2.6 GHz (5G)
The number of {train, test} data	{800, 200} per each (T_{IFFT}, T_{CP}, modulation) case
$\{\mathcal{S}_P, \mathcal{S}_Q\}$	{15, 50}
Time duration of each data	400 μs (Wi-Fi 6), 3.5 ms (5G)

4 EVALUATION

4.1 Evaluation environments

4.1.1 Data collection. The proposed classifier is evaluated with synthetic data generated from AWGN channel simulations with the details in Table 6. MATLAB R2023a WLAN and 5G toolbox [22] are deployed to generate the synthetic AWGN dataset. Wi-Fi HT [15] and HE format [14] are used to generate data with $T_{\text{IFFT}} = 3.2\,\mu s$ and $12.8\,\mu s$ in Wi-Fi 6. For 5G data, every SCS option in FR1, $\mu \in \{0, 1, 2\}$, is tested. All PHY channels listed in Table 4 are included in every 5G data item.

To evaluate whether the performance of the proposed system remains invariant across varying 5G PHY channel configurations, the parameters for allocating REs to PHY channels are set for each data type. For example in PDCCH, symbol duration, aggregation level, and starting symbol number are randomly selected. PHY broadcast channel (PBCH), primary synchronization signal (PSS), and secondary synchronization signal (SSS) are included only when $\mu \in \{0, 1\}$ since they are not available for $\mu = 2$. The other 5G PHY channel parameters are from FR1 test models in [1, 3].

4.1.2 Building classifier input. First, to avoid using the Wi-Fi preamble, we remove the first 2000 samples from each data. If the estimated T_{IFFT} corresponds to those of Wi-Fi 6, an IQ sequence whose length corresponds to 40+2 or 10+2 OFDM symbols is deployed to build the feature, $Y^i_f[k]$, starting with a random sample. We need an additional OFDM symbol due to the unknown starting index of an OFDM symbol sequence, $p \in [0, N_{\text{FFT}} - 1]$. Furthermore, one more extra OFDM symbol is required to evaluate phase differences between those of the last OFDM symbol and the next one. N_{null} is set to 8 and 32 for Wi-Fi HT and HE, respectively. If the estimated T_{IFFT} refers to 5G, $y'[n]$ of length (3 ms + 3 OFDM symbols) is used to estimate p and **firstIndexLongCP**. For 5G, the sequence of 14 OFDM symbols is utilized. We also evaluate the case using $Y^i_f[k]$ values as an input to assess how much the histogram input contributes to the performance. In this case, one data input consists of 2240 samples for Wi-Fi 6 or 7900 samples for 5G. The average number of feature elements in a single piece of 5G histogram data is 7858. We use fixed-duration data for a fair comparison, but the classifier can take the variable length data as input as the obtained feature is processed to a histogram using the

Table 7: SNR required for data communication with each modulation

Modulation	BPSK	QPSK	16QAM
SNR for Wi-Fi 6 (dB)	5	10	16
SNR for 5G (dB)	-	15	18
Modulation	64QAM	256QAM	1024QAM
SNR for Wi-Fi 6 (dB)	22	30	35
SNR for 5G (dB)	21	27	-

algorithms in Sec. 3.3. For both cases of input formats, an input with both phases of $\angle Y^i_{\Delta n}[k]$ modulo $\pi/2$ and $\angle Y^i_{\Delta n}[k]$ are evaluated.

4.2 Evaluation results

Results in Fig. 8 are obtained with synthetic AWGN channel data. Fig. 8a shows estimation accuracy of the OFDM parameters $\{T_{\text{CP}}, T_{\text{IFFT}}\}$. Normal CP and Extended CP in the legend refer to the shortest and longest option for T_{CP}, respectively, given Δf_{SCS}. Medium CP of Wi-Fi HE refers to 1.6 μs T_{CP}. In every case, accuracy is over 99%. In Fig. 8b, the estimation accuracy of *correctly finding* the starting index of an OFDM symbol is shown for the method in Sec. 3.2. *Correctly finding* means that the starting index time is within ϵ tolerance of the true time. In Fig. 8b, we note that the estimation accuracy for identifying the starting index of an OFDM symbol falls below 60% for both Wi-Fi 6 formats and 5G. When the tolerance is relaxed to $N_{\text{CP}}/4$ time samples, the reported estimation accuracy increases to 99%.

The accuracy of estimating an OFDM symbol with long CP is shown in Fig. 8c. Aside from $\Delta f_{\text{SCS}} = 60$ kHz, the performance is over 90% even at low SNR of 5 dB. Accuracy at $\Delta f_{\text{SCS}} = 60$ kHz is low because the duration of an OFDM symbol with long CP is larger than the others. The large number of symbols that the peak detection function needs to detect also negatively affects the peak detection performance. At $\Delta f_{\text{SCS}} = 60$ kHz, there are 30 peaks that should be identified in line 4 of Algorithm 1, which is considerably larger than the 9 or 16 peaks at $\Delta f_{\text{SCS}} = 15$ kHz and $\Delta f_{\text{SCS}} = 30$ kHz.

Figure 9 shows modulation classification accuracy with AWGN channel data. The proposed algorithm with a histogram input with the phases $\Delta\angle Y^i_{\Delta n}[k]$ modulo $\pi/2$ outperforms in all considered cases, except for Wi-Fi 6 at 5 dB SNR. The performance gap between using the histogram as classifier input as opposed to using the feature value input increases in Wi-Fi HE and even more so in 5G. This is because the histogram input helps the classifier to discriminate the detailed symbol constellation of high-order modulations.

In Table 8, the accuracy of each modulation format is shown when the SNR satisfies the minimum requirement for standard-compliant data communication. We deploy error vector magnitude (EVM) levels required for data communication with each modulation for Wi-Fi 6 and 5G documentations [1, 14]. Required SNR values are calculated using the relation between EVM and SNR presented in [26]. SNR values required for the smallest coding rate are chosen for each modulation and chosen values are arranged in Table 7. For every modulation with both Wi-Fi 6 formats and 5G, accuracy is at least 98%.

Table 8: Accuracy when SNR is over the minimum requirements for standard-compliant data communication

Modulation	BPSK	QPSK	16QAM
Wi-Fi HT	100%	100%	100%
Wi-Fi HE	100%	100%	100%
5G	-	99%	100%
Modulation	64QAM	256QAM	1024QAM
Wi-Fi HT	100%	-	-
Wi-Fi HE	100%	100%	98%
5G	100%	100%	-

5 CONCLUSION

Modulation classification of Wi-Fi 6 and 5G signals for spectrum sensing is studied. Our system deploys CAF to estimate SCS and CP length and achieve 99% accuracy. Without control information, our proposed preprocessing algorithm extracts features characterizing modulation schemes insensitive to synchronization errors. The preprocessing stage also estimates the CP position and the symbol with long CP of 5G signals. The form of the features is converted to be more suitable as inputs for the CNN-based classifier, which contributes to performance improvement in identifying high-order modulation. With data under various protocol configurations, our system identifies modulations of OFDM signals with 98% classification accuracy when SNR is higher than the value required for data transmission given a modulation. We are planning to extend this study to multiple-input multiple-output (MIMO) and orthogonal frequency division multiple access (OFDMA) scenario, to make more general transmission cases covered.

REFERENCES

[1] 3GPP TR 38.141-1. 2023. NR; Radio Resource Control (RRC); Base Station (BS) conformance testing; Part 1: Conducted conformance testing. ver. 18.1.0.

[2] 3GPP TR 38.331. 2023. NR; Radio Resource Control (RRC); Protocol specification. ver 17.4.0.

[3] 3GPP TR 38.521-4. 2023. NR; User Equipment (UE) conformance specification; Radio transmission and reception; Part 4: Performance requirements. ver 17.2.1.

[4] Dhamyaa Husam Al-Nuaimi, Nor Ashidi Mat Isa, Muhammad Firdaus Akbar, and Intan Sorfina Zainal Abidin. 2021. AMC2-Pyramid: Intelligent pyramidal feature engineering and multi-Distance decision making for automatic multi-carrier modulation classification. *IEEE Access* 9 (Sept. 2021), 137560–137583. https://doi.org/10.1109/ACCESS.2021.3115888

[5] Nicola Bui and Joerg Widmer. 2016. OWL: A reliable online watcher for LTE control channel measurements. In *Proc. 5th Workshop on All Things Cellular: Operations, Applications and Challenges*. Association for Computing Machinery, New York (NY), USA, 25–30. https://doi.org/10.1145/2980055.2980057

[6] Robert Falkenberg and Christian Wietfeld. 2019. FALCON: An accurate real-time monitor for client-based mobile network data analytics. In *Proc. IEEE GLOBECOM*. IEEE, Waikoloa (HI), USA, 1–7. https://doi.org/10.1109/GLOBECOM38437.2019.9014096

[7] Jiabao Gao, Xuemei Yi, Caijun Zhong, Xiaoming Chen, and Zhaoyang Zhang. 2019. Deep learning for spectrum sensing. *IEEE Wireless Commun. Letters* 8, 6 (2019), 1727–1730. https://doi.org/10.1109/LWC.2019.2939314

[8] William A Gardner and Chad M Spooner. 1994. The cumulant theory of cyclostationary time-series. I. Foundation. *IEEE Trans. Signal Process.* 42, 12 (Dec. 1994), 3387–3408. https://doi.org/10.1109/78.340775

[9] Rahul Gupta, Sushant Kumar, and Sudhan Majhi. 2020. Blind modulation classification for asynchronous OFDM systems over unknown signal parameters and channel statistics. *IEEE Trans. Veh. Technol.* 69, 5 (Mar. 2020), 5281–5292. https://doi.org/10.1109/TVT.2020.2981935

[10] Tuan Dinh Hoang, CheolJun Park, Mincheol Son, Taekkyung Oh, Sangwook Bae, Junho Ahn, Beomseok Oh, and Yongdae Kim. 2023. LTESniffer: An open-source LTE downlink/uplink eavesdropper. In *Proc. 16th ACM Conference on Security*

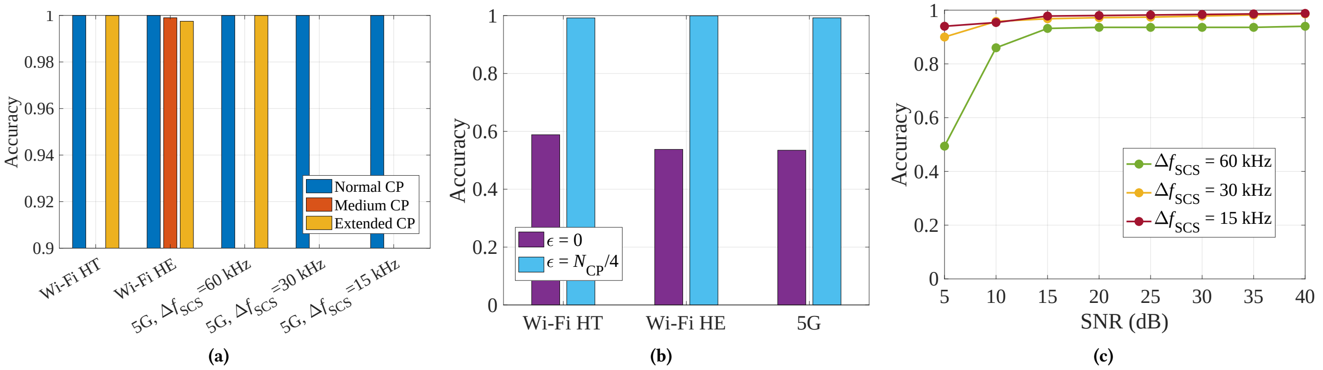


Figure 8: OFDM parameter estimation results: (a) Accuracy for estimating T_{IFFT} and T_{CP}, (b) Accuracy for choosing the first index of CP with acceptable error ϵ, and (c) Accuracy for finding an OFDM symbol with long CP of 5G signals.

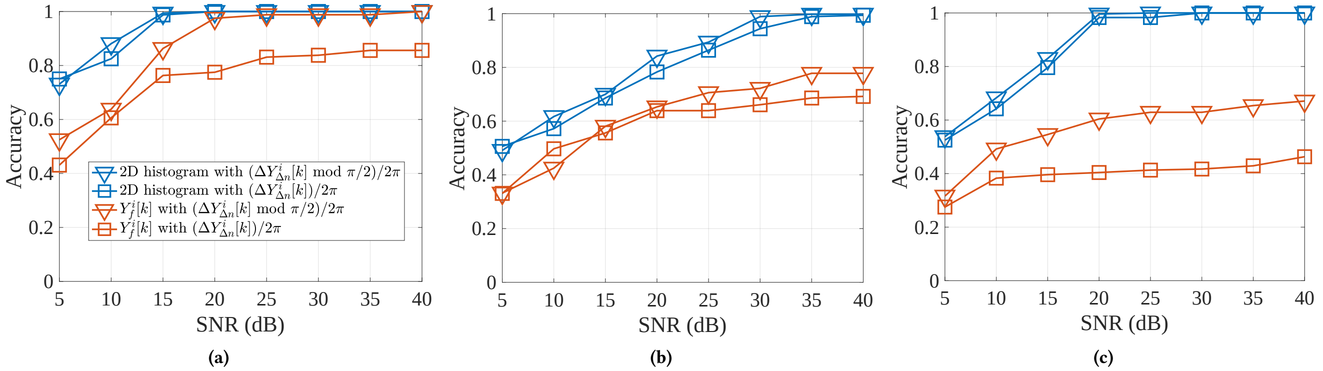


Figure 9: Classification accuracy for modulations vs. SNR: (a) Wi-Fi HT, (b) Wi-Fi HE, (c) 5G.

and Privacy in Wireless and Mobile Networks (WiSec). Guildford, UK, 43—-48. https://doi.org/10.1145/3558482.3590196

[11] Sheng Hong, Yu Wang, Yuwen Pan, Hao Gu, Miao Liu, Jie Yang, and Guan Gui. 2020. Convolutional neural network aided signal modulation recognition in OFDM systems. In *Proc. IEEE VTC*. 1–5. https://doi.org/10.1109/VTC2020-Spring48590.2020.9128455

[12] Sheng Hong, Yibin Zhang, Yu Wang, Hao Gu, Guan Gui, and Hikmet Sari. 2019. Deep learning-based signal modulation identification in OFDM systems. *IEEE Access* 7 (Aug. 2019), 114631–114638. https://doi.org/10.1109/ACCESS.2021.3102223

[13] Steven Siying Hong and Sachin Rajsekhar Katti. 2011. DOF: A local wireless information plane. In *Proc. ACM SIGCOMM*. 230–241. https://doi.org/10.1145/2018436.2018463

[14] IEEE 802.11ax. 2021. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 1: enhancements for high-efficiency WLAN.

[15] IEEE 802.11n. 2009. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 5: enhancements for higher throughput.

[16] Anand Kumar, Keerthi Kumar Srinivas, and Sudhan Majhi. 2023. Automatic Modulation Classification for Adaptive OFDM Systems Using Convolutional Neural Networks with Residual Learning. *IEEE Access* 11 (Jun. 2023), 61013–61024. https://doi.org/10.1109/ACCESS.2023.3286939

[17] Swarun Kumar, Ezzeldin Hamed, Dina Katabi, and Li Erran Li. 2014. LTE radio analytics made easy and accessible. *Proc. ACM SIGCOMM* 44, 4 (Aug. 2014), 211–222. https://doi.org/10.1145/2740070.2626320

[18] Yan Li, Johan Barthelemy, Shuai Sun, Pascal Perez, and Bill Moran. 2020. A case study of WiFi sniffing performance evaluation. *IEEE Access* 8 (2020), 129224–129235. https://doi.org/10.1109/ACCESS.2020.3008533

[19] Dancheng Liu, Kazim Ergun, and Tajana Šimunić Rosing. 2023. Towards a Robust and Efficient Classifier for Real World Radio Signal Modulation Classification. In *Proc. IEEE ICASSP*. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10094907

[20] Golledge Electronics Ltd. 2023. *GTXO-203T | 1.8V~3.6V SM TCXO | Golledge*. Retrieved 06.21.2023 from https://www.golledge.com/products/gtxo-203t-ultra-miniature-tight-stability-tcxo/c-26/p-287/

[21] Norbert Ludant, Pieter Robyns, and Guevara Noubir. 2023. From 5G Sniffing to Harvesting Leakages of Privacy-Preserving Messengers. In *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society, Los Alamitos, CA, USA, 1919–1934. https://doi.org/10.1109/SP46215.2023.00110

[22] MathWorks. 2023. *MATLAB Products*. Retrieved 06.21.2023 from https://www.mathworks.com/products.html/

[23] Myung Chul Park and Dong Seog Han. 2021. Deep learning-based automatic modulation classification with blind OFDM parameter estimation. *IEEE Access* 9 (2021), 108305–108317. https://doi.org/10.1109/ACCESS.2021.3102223

[24] Amit Kumar Pathy, Anand Kumar, Rahul Gupta, Sushant Kumar, and Sudhan Majhi. 2021. Design and implementation of blind modulation classification for asynchronous MIMO-OFDM system. *IEEE Trans. Instrum. Meas.* 70 (Sept. 2021), 1–11. https://doi.org/10.1109/TIM.2021.3109737

[25] Anjana Punchihewa, Vijay K Bhargava, and Charles Despins. 2011. Blind estimation of OFDM parameters in cognitive radio networks. *IEEE Trans. Wireless Commun.* 10, 3 (Mar. 2011), 733–738. https://doi.org/10.1109/TWC.2010.010411.100276

[26] Rishad Ahmed Shafik, Md Shahriar Rahman, and AHM Razibul Islam. 2006. On the extended relationships among EVM, BER and SNR as performance metrics. In *Proc. IEEE ICECE*. 408–411. https://doi.org/10.1109/ICECE.2006.355657

[27] Ling Yu, Jin Chen, and Guoru Ding. 2017. Spectrum prediction via long short term memory. In *Proc. IEEE ICCC*. 643–647.

[28] Zufan Zhang, Hao Luo, Chun Wang, Chenquan Gan, and Yong Xiang. 2020. Automatic modulation classification using CNN-LSTM based dual-stream structure. *IEEE Trans. Veh. Technol.* 69, 11 (Nov. 2020), 13521–13531.