PhaseAnti: An Anti-Interference WiFi-Based Activity Recognition System Using Interference-Independent Phase Component

Jinyang Huang, Bin Liu, Chenglin Miao, *Member, IEEE*, Yan Lu, Qijia Zheng, Yu Wu, Jiancun Liu, Lu Su, *Member, IEEE*, and Chang Wen Chen, *Fellow, IEEE*

Abstract—Driven by a wide range of essential applications, significant achievements have recently been made to explore WiFi-based Human Activity Recognition (HAR) techniques that utilize the information collected by commercial off-the-shelf (COTS) WiFi infrastructures to infer human activities without the need for the subject to carry any devices. Although existing WiFi-based HAR systems achieve satisfactory performance in some instances, they are faced with a severe challenge that the impacts of ubiquitous Co-channel Interference (CCI) on WiFi signals are inevitable. This downgrades the performance of these HAR systems significantly. To address this challenge, we propose PhaseAnti in this paper, a novel WiFi-based HAR system to exploit the CCI-independent phase component, Nonlinear Phase Error Variation (NLPEV), of WiFi Channel State Information (CSI) to cope with the negative effects of CCI. The stability of NLPEV data and the sensibility of this component to motions are rigorously analyzed. Furthermore, validated by extensive properly designed experiments, this phase component across subcarriers is invariant under various CCI scenarios while sufficiently distinct for different motions. Therefore, the NLPEV data can be used and processed effectively to perform HAR in CCI scenarios. Extensive experiments with various daily activities in different indoor rooms demonstrate the superior effectiveness and generalizability of the proposed PhaseAnti system under various CCI scenarios. Specifically, PhaseAnti achieves a 96.5% recognition accuracy rate (RAR) on average in different CCI scenarios, which can improve up to a 16.7% RAR compared with the amplitude component in the presence of CCI. Furthermore, the recognition speed is 10.3 $\times$ faster than the state-of-the-art solution.

Index Terms—Activity recognition, WiFi-based, co-channel interference, interference-independent component

1 Introduction

With the widespread deployments of the Internet of Things (IoT), Human Activity Recognition (HAR) has become a crucial service in many IoT applications, such as smart cities, smart homes, and healthcare [1]. Significant attention has been received from both academia and industry, with diverse solutions based on cameras [2] and wearable sensors [3]. However, camera-based approaches have the fundamental limitations of requiring enough light, and the privacy leak problem cannot be ignored. Wearable sensor-based approaches are sometimes inconvenient and uncomfortable due to the sensors that subjects have to wear on their bodies. Recently, WiFi-based HAR systems [4], [5], [6], [7] have been proposed since different human activities cause distinct multipath distortions in WiFi signals. Their key advantages over the camera and wearable sensor-based systems are that they do not need light, preserve user privacy, and do not require subjects to carry any devices as they rely on the signals reflected by the human body.

However, the crucial limitation of WiFi-based HAR systems lies in ignoring the impact of Co-channel Interference (CCI) on WiFi signals, which results in a sharp drop in their system performance under CCI scenarios. Unfortunately, CCI becomes more ubiquitous due to the proliferating of the number and types of WiFi devices in the last decade and often varies due to the channel hopping mechanism [8]. Thus, it is almost impossible to find a channel that is clean or only occupied by one device within the signal range. As shown in Fig. 1, the mobile phone receives multiple router signals in one place, but the spectrum of some router channels is overlapped with the connection channel spectrum. Thus, CCI happens. Moreover, CCI has severe negative effects on WiFi signals, e.g., the reduction of received packet number [8], the confusion of signal amplitude [9], and the weakening of subcarrier correlation [10], which further results in the degradation of WiFi HAR system performance. Therefore, an anti-CCI WiFi-based HAR system to recognize motions accurately in CCI scenarios is necessary and valuable.

Indeed, for an anti-CCI HAR system, the signal component used for recognition should be invariant across various

- *Jinyang Huang, Bin Liu, Yan Lu, Qijia Zheng, Yu Wu, and Jiancun Liu are with the CAS Key Laboratory of Electromagnetic Space Information, University of Science and Technology of China, Hefei 230026, China. E-mail: {huangjy, luyan17, zhengqj, wy666, jsen617}@mail.ustc.edu.cn, flowice@ustc.edu.cn.*
- *Chenglin Miao is with the Department of Computer Science, University of Georgia, Athens, GA 30602 USA. E-mail: cmiao@uga.edu.*
- *Lu Su is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA. E-mail: lusu@buffalo.edu.*
- *Chang Wen Chen is with the School of Science and Engineering, Hong Kong Polytechnic University, Hong Kong 999077, China. E-mail: changwen.chen@polyu.edu.hk.*

Manuscript received 16 July 2021; revised 6 Oct. 2021; accepted 3 Nov. 2021. Date of publication 12 Nov. 2021; date of current version 4 Apr. 2023. (Corresponding author: Bin Liu.)
Digital Object Identifier no. 10.1109/TMC.2021.3127721

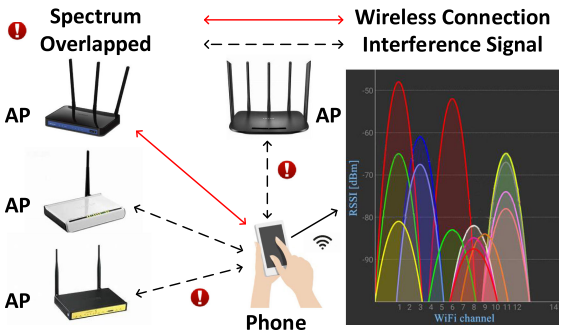


Fig. 1. Indoor environment of electromagnetic.

CCI scenarios and time but sensitive and distinct for different motions. If the signal component itself changes with CCI, it is difficult to extract the only motion-related part from this signal component. However, the Channel State Information (CSI) amplitude used by most WiFi-based HAR systems [5], [11], [12] is varying in CCI scenarios since different devices within the signal coverage adjust their transmitter power to better compete for the channel [10]. Thus, CSI amplitudes do not satisfy the requirement of the anti-CCI signal component. Fortunately, CSI phases are not affected by the varying transmission power caused by CCI. Therefore, the CCI-independent component can be extracted from the CSI phase.

Three challenges need to be formally addressed before realizing a novel anti-CCI CSI-based HAR system.

- *Varying CSIs:* The CSI amplitude used by most state-of-the-art HAR systems is varying when CCI changes. In addition, due to the significant variations caused by noises and the unsynchronized time and frequency at transmitter and receiver, the CSI phase initially contains many errors [13], and these phase errors are difficult to eliminate. Therefore, it is challenging to extract the CCI robustness and only motion-related signal component from varying CSIs.
- *Based on universal equipment:* The specialized radio frequency equipment can select a unique channel to avoid CCI, and the impact of CCI can also be degraded by boosting the transmitter power. However, these methods are expensive to implement in daily life or have side effects on other wireless signals, e.g., Bluetooth and Zigbee. Thus, how to realize the complex function of anti-CCI with universal equipment becomes a question worthy of consideration.
- *Varying CCI scenarios:* The impact levels of various CCI scenarios on CSIs are distinct. Therefore, a fixed scheme is unsuitable for all cases since the tradeoff between accuracy and recognition speed in different CCI scenarios is distinct. Apparently, how to design a flexible scheme to deal with different CCI scenarios validly becomes an inevitable problem.

To tackle these challenges, we propose PhaseAnti, an anti-CCI HAR system based on WiFi CSI. Specifically, in CCI-scenarios, by eliminating irrelevant errors, the phase component Nonlinear Phase Error Variation (NLPEV) is leveraged from the commercial off-the-shelf (COTS) WiFi device since this component keeps constant for various CCI scenarios and contains motion information. By using this extracted phase data to perform HAR in CCI scenarios, the activity can be recognized accurately with a low recognition time. On the contrary, in non-interfering scenarios, the amplitude data are exploited to further improve the recognition speed.

In total, we make the following contributions:

- We propose a novel Anti-interference, Non-intrusive HAR system PhaseAnti leveraging CSI from a single COTS WiFi device. As far as we know, this work is the first to present a CCI-independent component of WiFi signals and the first to exploit the CSI phase component to perform HAR in CCI scenarios. We carefully verify the invariance of the proposed component NLPEV to various CCI scenarios and the difference of this component to different activities.
- Instead of simply using a fixed component, we design a flexible scheme to adaptively choose different WiFi CSI components in different CCI scenarios to better balance the recognition accuracy rate (RAR) and the recognition speed of the system. Specifically, the NLPEV is exploited in CCI scenarios to improve the system performance, and the amplitude is used in non-interfering scenarios for further boosting the recognition speed.
- Extensive experiments with different activities have been performed in various CCI scenarios. The results show that PhaseAnti improves up to 19.7% RAR on average in the presence of complex CCI, reaching 96.9%, and the recognition speed is $10.3\times$ faster than the pioneer anti-CCI HAR solution [10].

The rest of the paper is organized as follows. Section 2 discusses the related works. Then, we present the preliminaries of CSI and the analysis of NLPEV data in Section 3. Next, we describe the PhaseAnti system design in Section 4. Implementation, evaluation, and the impacts of various factors on PhaseAnti performance are presented in Section 5. Finally, we conclude our work in Section 6.

2 Related Work

According to the signal employed for motion detection, existing device-free wireless sensing systems can be broadly classified into three categories: Dedicated Radio Frequency (DRF)-based, Received Signal Strength Indication (RSSI)-based, and CSI-based. Specifically, there are three requirements for these HAR systems, including informative measurements, universal, and robustness to CCI.

- *Informative measurements*: Contain information to recognize different activities accurately.
- *Universal*: Use existing equipment or deployable on COTS infrastructures and do not affect other functions.
- *Robustness to interference*: Mitigate or eliminate the effects of CCI.

To summarize the pros and cons of different wireless sensing HAR methods, a Venn diagram with the three features of informative measurements, universal, and robustness to CCI is drawn in Fig. 2. This diagram is employed to differentiate various approaches and shows which solution can deal with which requirements.

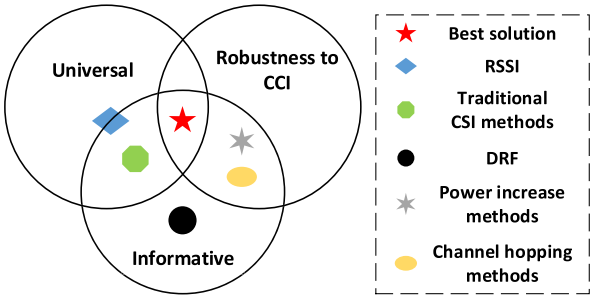


Fig. 2. Difference among various wireless signal-based HAR systems.

DRF-Based. Fine-grained radio frequency signals can be collected by the dedicatedly designed hardware [14], [15], [16]. WiSee [14] used Universal Software Radio Peripheral to capture the Doppler shift in body reflected signals to recognize a set of nine gestures with high accuracy. TagFree [16] employed the Thingmagic reader to collect the multipath signals from different RFID tags to classify the working motions. Although the RARs of these systems are usually satisfactory, the specialized equipment used by these methods is irreplaceable but expensive. Thus, DRF-based systems are expressed as a black circle in Fig. 2.

RSSI-Based. RSSI is an indication to measure the power of received radio signals. Since different activities cause distinct RSSI fluctuations, activities can be recognized accordingly by effective signal processing. PAWS [6] explored WiFi ambient signals to establish RSSI fingerprints of different activities. Wigest [7] leveraged RSSI changes to sense in-air hand gestures around the mobile device. However, the RSSI used by these methods falls entirely in the time domain, while the frequency domain is totally neglected. Suffering from performance degradation due to the multipath effect is also a problem for RSSI-based systems. Therefore, a blue diamond in Fig. 2 can clearly represent RSSI-based HAR systems.

CSI-Based. In contrast, CSI is a fine-grained value derived from the physical layer of the COTS WiFi device. Features are descriptions of activities from different perspectives, i.e., time domain and frequency domain. Thus, compared with RSSI, more information can be obtained from CSIs.

Many good solutions have been proposed. E-eyes [17] presented location-oriented activity identification, such as cooking and gaming. The profile can be adaptively updated by E-eyes to accommodate the movement or replacement of wireless devices, and thus the day-to-day motion profile was calibrated. Considering that traditional methods are based on similarity classification to achieve recognition, CARM [18] first quantified the correlation between CSI value dynamics and human movement speeds as well as the correlation between the movement speeds of different human body parts and a specific human activity. By this way, CARM quantitatively built the correlation between CSI value dynamics and a specific human activity for the more robust recognition of motions. FinderDraw [19] proposed a CSI-quotient model leveraging two antennas of a WiFi receiver to achieve finger motion sensing. However, the metal plate with the specific dielectric coefficient is necessary for FinderDraw to act as a perfect signal reflector. The Convolutional Neural Network was used by SignFi [20] as the classification algorithm to process WiFi packets for the sign language gesture recognition. Widar 3.0 [21] developed a one-fits-all deep learning model to fully exploit spatial-temporal characteristics of the proposed domain-independent feature for gesture recognition. Compared with machine learning methods, these two deep learning classifiers achieve more satisfactory performance but cost more recognition time. FallDeFi [22] employed the conventional Short-Time Fourier Transform to extract time-frequency features and a sequential forward selection algorithm to single out features that are resilient to environmental changes to further achieve a high fall detection rate. The relation between Doppler shifts and motion directions was calculated by WiDance [23] to achieve the human-computer interaction on the dance-pad. PhaseBeat [24] exploited the CSI phase difference to monitor breathing and heartbeats. However, these existing works based on WiFi do not consider the impacts of CCI caused by other WiFi devices, and their experimental results are based on a non-CCI environment. Furthermore, the performance of these systems degrades due to CCI. Based on these reasons, these traditional CSI-based HAR systems are expressed as a green octagon.

The specialized radio frequency equipment can select a unique channel to avoid CCI (represented by a yellow oval), and the impact of CCI can also be degraded by increasing the transmitter power (expressed by a gray hexagon). However, since these approaches are expensive to implement in daily life or have side effects on other wireless signals, e.g., Bluetooth and Zigbee, these methods do not satisfy the universal requirement. Although WiAnti [10] proposed a subcarrier selection algorithm to select the most informative subcarriers to realize anti-interference, this algorithm still stays at the signal processing level and does not propose the CCI-independent CSI component. Furthermore, the varying CSI amplitude caused by CCI limits its performance.

To overcome these drawbacks, we present PhaseAnti, a CSI-based HAR system that is robust to CCI. Inspired by the prior works, our work's contribution lies in obtaining the CCI-independent CSI component NLPEV. To the best of our knowledge, it is the first to leverage the CCI-independent CSI component from a COTS WiFi device to recognize activities accurately in CCI scenarios. This means the proposed system PhaseAnti satisfies all three requirements.

3 Preliminaries and Observations

3.1 Overview of Channel State Information

Network interface cards (NICs) continuously capture variations in the wireless channel using CSI, which is fine-grained physical layer information and characterizes the Channel Frequency Response (CFR) of the wireless channel [8]. CSI reveals the channel characteristics experienced by the received signals, such as the effect of scattering, multipath effect, and power decay [9]. Since WiFi systems commonly apply *Orthogonal Frequency Division Multiplexing* (OFDM) technology, the channel between each transmitter-receiver (Tx-Rx) antenna pair consists of multiple subcarriers [11]. Let $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ be the frequency domain representations of the transmitted and the received signals of the i^{th} Tx-Rx pair, respectively. Hence, the two signals can be related by the expression

$$\boldsymbol{Y}_i = \boldsymbol{H}_i \boldsymbol{X}_i + \boldsymbol{\sigma}_i, \tag{1}$$

where $\boldsymbol{H}_i$ is the complex-valued CFR of the i^{th} Tx-Rx pair, which can be estimated by transmitting a known preamble of OFDM symbols between the transmitter and the receiver [11], and $\boldsymbol{\sigma}_i$ is the additive white Gaussian noise [24].

As the IEEE 802.11n standard [8] is a protocol based on OFDM technology, the 2.4 to 2.4835GHz band is divided into 14 channels, where each channel has 56 subcarriers in 20M bandwidth. Therefore, the estimated value of $\boldsymbol{H}_i$ for 56 subcarriers can be represented as

$$\boldsymbol{H}_i = [\boldsymbol{h}_i{}^{(1)}, \boldsymbol{h}_i{}^{(2)}, \ldots, \boldsymbol{h}_i{}^{(k)}, \ldots \boldsymbol{h}_i{}^{(56)}], \tag{2}$$

where $\boldsymbol{h}_i{}^{(k)}$ is complex-valued CFR of the k^{th} subcarrier in i^{th} Tx-Rx pair. CSI measurements basically contain these CFR values, and $\boldsymbol{h}_i{}^{(k)}$ can be expressed as

$$\boldsymbol{h}_i{}^{(k)} = \left|\boldsymbol{h}_i{}^{(k)}\right| \cdot e^{-j \cdot \angle \boldsymbol{h}_i{}^{(k)}} = I_i{}^{(k)} + jQ_i{}^{(k)} \quad k \in \boldsymbol{K}, \tag{3}$$

where $\left|\boldsymbol{h}_i{}^{(k)}\right|$ and $\angle \boldsymbol{h}_i{}^{(k)}$ denote the amplitude and the phase of k^{th} subcarrier in i^{th} Tx-Rx pair. Besides, the raw CFRs estimated in NICs can also be recorded as the I/Q signal. I and Q are the in-phase component and the quadrature component, respectively. $\boldsymbol{K}$ contains the subcarrier indexes. Although the WiFi system has 56 subcarriers over a 20 MHz channel, the Intel 5300 NIC we use can only report CSI for 30 of the 56 subcarriers [4]. Specifically, for the Intel 5300 NIC, $\boldsymbol{K} = [-28, -26, \ldots, -2, -1, 1, 3, \ldots, 27, 28]$. The phase $\angle \boldsymbol{h}_i{}^{(k)}$ ($\boldsymbol{\phi}$) can be calculated from the I/Q component

$$\boldsymbol{\phi} = \arctan\left(\frac{Q}{I}\right). \tag{4}$$

3.2 Effect of Motion on CSI

Since human bodies and surrounding objects reflect radio signals, a transmitted signal arrives at the receiver through multiple paths. If a wireless signal (k^{th} subcarrier in i^{th} Tx-Rx pair) arrives at the receiver through Υ different paths, $\boldsymbol{h}_i{}^{(k)}(\vartheta, t)$ can be given by the following equation [18], [25]

$$\boldsymbol{h}_i{}^{(k)}(\vartheta, t) = e^{-j \cdot 2\pi \Delta \vartheta t} \sum_{\wp=1}^{\Upsilon} \varpi_{\wp}(\vartheta, t) \cdot e^{-j \cdot 2\pi \vartheta \tau_{\wp}(t)}, \tag{5}$$

where j is the imaginary unit, ϑ denotes the k^{th} subcarrier frequency, $\varpi_{\wp}(\vartheta, t)$ denotes the complex-valued representation of attenuation and initial phase offset of the $\wp^{th}$ path, and $e^{-j \cdot 2\pi \vartheta \tau_{\wp}(t)}$ is the phase shift on the $\wp^{th}$ path which has a propagation delay of $\tau_{\wp}(t)$. In addition, $e^{-j \cdot 2\pi \Delta \vartheta t}$ is the phase shift caused by the subcarrier frequency difference between the transmitter and the receiver.

The movements indeed cause CSI fluctuations by changing the electromagnetic distribution in the environment. According to [13], [18], [25], the path length changing by one wavelength causes the receiver to experience a phase shift of 2π in the corresponding subcarrier. Furthermore, since human movements change different path phases at the receiver, the different paths are superimposed according to the new phase relationship [9], which finally results in the fluctuations of subcarrier amplitudes. Thus, human activities make both phases and amplitudes fluctuate. Additionally, PhaseBeat [24] demonstrates that the true phase at the receiver is a periodic signal for the same frequency with the periodic activity. Therefore, the movement frequency can be captured by the corresponding CSI frequency responses.

3.3 Impact of CCI on CSI

The negative impacts of CCI on CSI significantly downgrade the recognition performance of the WiFi-based HAR systems. Specifically, three main negative effects are introduced as follows.

First, the sampling rate is reduced by CCI. According to the *Carrier Sense Multiple Access with Collision Avoidance* mechanism, a node can send packets only when it detects that the working channel is idle. The waste of time is inevitable for the nodes to wait for the channel idle, which eventually leads to the number reduction of the packets received per unit time. Therefore, the sampling rate of the HAR system is reduced accordingly, and the motion capture ability of the WiFi-based HAR systems downgrades.

Second, CCI weakens the subcarrier correlation, which leads to the data dimensionality reduction algorithms of the state-of-the-art HAR systems invalid. This is because the data dimensionality reduction algorithms of most existing WiFi-based HAR systems [5], [10] rely on the strong correlation of subcarriers. However, to cope with CCI, different nodes within the range increase the energy of subcarriers in the overlapping area, which leads to the energy imbalance between subcarriers and weakens the subcarrier correlation. The weak subcarrier correlation makes these data dimensionality reduction algorithms ineffective and ultimately leads to the performance degradation of these WiFi-based HAR systems.

Third, CCI makes the difference between different samples of the same motion under different CCI conditions become significant, which affects the normal classification of the activity. Since different CCI conditions in the varying CCI scenario lead to different subcarrier energy distributions, for the same activity, the power difference between the different motion samples under different CCI conditions becomes larger, which directly leads to the unstableness of the activity features and further downgrades the recognition performance of the WiFi-based HAR systems.

CCI is very universal, and it can significantly affect the performance of the existing WiFi-based HAR systems. Therefore, it is necessary to propose a CCI robust WiFi-based HAR system for accurately recognizing human activities in various CCI scenarios.

3.4 NLPEV

Due to the unstable signal transmission environment and imperfect hardware design, the receiver's measured phases are different from the phases at the transmitter. Generally, the phase difference between transmitter and receiver can be grouped into two categories, i.e., linear phase error (LPE) and nonlinear phase error (NLPE) [26]. Particularly, LPE and NLPE mean that the phase errors vary linearly and nonlinearly with subcarrier indexes, respectively.

According to [13], [25], [27], for a particular pair of transmitter and receiver, the subcarrier phases $\boldsymbol{\phi}$ measured at the receiver can be expressed as

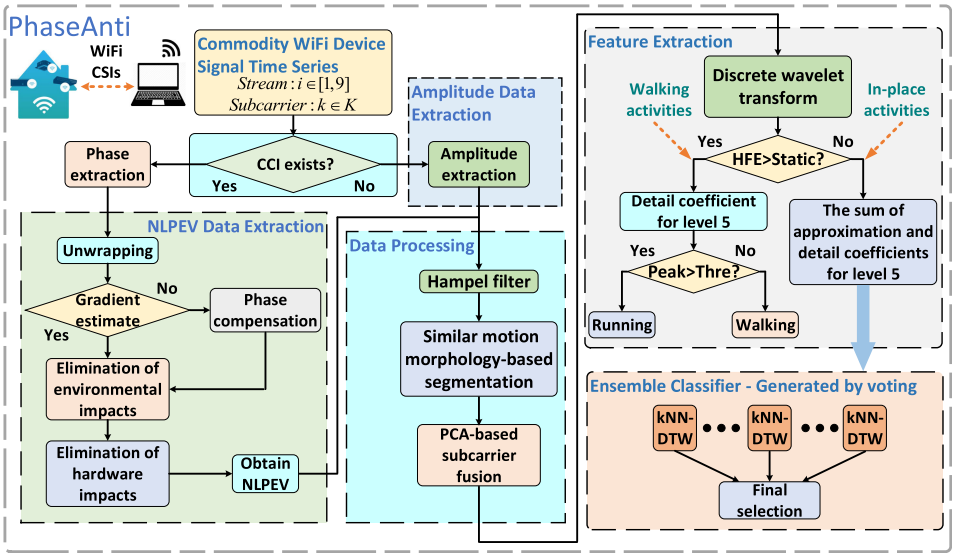


Fig. 3. PhaseAnti workflow.

$$\boldsymbol{\phi} = \boldsymbol{\varphi}_{re} + \boldsymbol{L}_{pbd} + \boldsymbol{L}_{sfo} + \boldsymbol{L}_{cfo} + \boldsymbol{L}_{tof} + \boldsymbol{N}_{to}, \tag{6}$$

where $\boldsymbol{\varphi}_{re}$ represents the real phases at the transmitter. $\boldsymbol{L}_{pbd}$, $\boldsymbol{L}_{sfo}$, $\boldsymbol{L}_{cfo}$, and $\boldsymbol{L}_{tof}$ denote the phase offsets due to packet boundary detection (PBD), sampling frequency offset (SFO), carrier frequency offset (CFO), and time of flight (ToF) respectively, and these phase offsets are LPEs [27]. According to [26], imperfect hardware design causes an NLPE. Besides, by affecting multipath, the activity also causes an NLPE [18]. Thus, the last element $\boldsymbol{N}_{to}$ can be written as

$$\boldsymbol{N}_{to} = \boldsymbol{N}_{ha} + \boldsymbol{N}_{mo}, \tag{7}$$

where $\boldsymbol{N}_{ha}$ and $\boldsymbol{N}_{mo}$ are the NLPEs caused by imperfect hardware design and human motion, respectively. Since $\boldsymbol{N}_{ha}$ is proved to be a constant for each specific NIC [26], the change of $\boldsymbol{N}_{to}$ equals the variation of $\boldsymbol{N}_{mo}$ caused by human movements. Therefore, by comparing the $\boldsymbol{N}_{to}$ in the movement with the $\boldsymbol{N}_{to}^{em}$ in an empty room without the human motion, the NLPE $\boldsymbol{N}_{mo}$ caused by human movements can be estimated as

$$\begin{aligned} \boldsymbol{N}_{to} - \boldsymbol{N}_{to}^{em} &= \boldsymbol{N}_{ha} + \boldsymbol{N}_{mo} - (\boldsymbol{N}_{ha} + \boldsymbol{N}_{mo}^{em}) \\ \boldsymbol{N}_{mo} &= \boldsymbol{N}_{to} - \boldsymbol{N}_{to}^{em}, \end{aligned} \tag{8}$$

where $\boldsymbol{N}_{mo}^{em}$ is the variation of NLPE without the human motion and equals to 0 [26]. Thus, $\boldsymbol{N}_{mo}$ can be estimated by Nonlinear Phase Error Variation (NLPEV) when NICs are fixed. Section 4 shows the specific derivation process of $\boldsymbol{N}_{mo}$. The NLPEV means the variation of the nonlinear phase error. In particular, most authoritative related works [24], [26], [27], [28] define the phase difference between the phases $\boldsymbol{\phi}$ measured at the receiver and the real phases $\boldsymbol{\varphi}_{re}$ at the transmitter as the phase errors. Since the proposed component is the variation of the nonlinear part of these phase errors, we call the proposed component as Nonlinear Phase Error Variation (NLPEV).

4 System Design and Methodology

4.1 System Overview

To make the WiFi-based HAR system robust to CCI, we design PhaseAnti, and its framework is shown in Fig. 3. First, we consider the classification of ten common activities or states. Specifically, nine different activities (standing, walking, running, sitting, push-ups, lift dumbbells, squats, stoop-down, and sit-ups) and a reference state (the empty room) are designed to observe the impacts of CCI on CSI signals with or without human movement. The sketch of each motion is shown in Fig. 4. The activities are divided into two categories, i.e., in-place activities (IPAs) and walking activities (WAs), according to whether they contain location changes. The IPAs are standing, sitting, push-ups, lift dumbbells, squats, stoop-down, and sit-ups. The WAs are walking and running. The walking and running motions are the shuttle walk and the shuttle run, respectively. It is worth mentioning that the PhaseAnti system can be easily extended according to specific motions in other real-life applications.

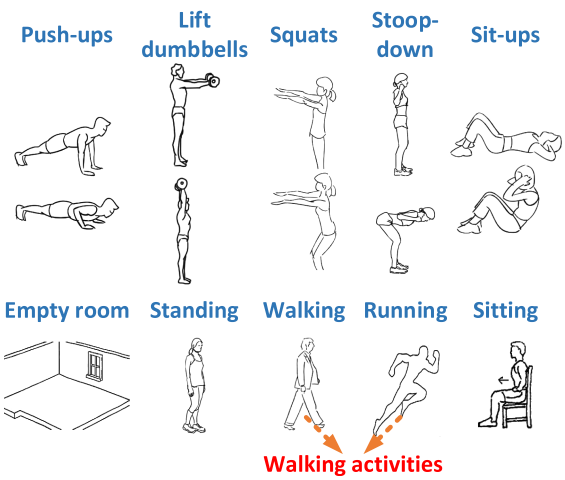


Fig. 4. Activity set: nine regular motions and one reference state.

The PhaseAnti system workflow mainly contains four parts: Judgment of CCI, Adaptive Data Processing, Feature Extraction, and Classification. Since different CCI states, e.g., non-interfering, and CCI existing, have distinct impacts on CSI, to better balance the recognition performance and the recognition speed of the system, the CCI state is first detected. The CCI existence is determined by recording the number of received frames per unit time due to the fact that the existence of CCI causes the number of received frames per unit time to decrease [10].

Next, according to the judgment result, different signal components are extracted and processed accordingly to cope with different CCI scenarios. Specifically, in CCI scenarios, the NLPEV information is extracted to mitigate the impacts of CCI on CSI since this component is CCI-independent. On the contrary, in non-interfering scenarios, the amplitude information is employed to further reduce the calculation complexity. After the HAR CSI component is determined and extracted, the same data processing algorithms, i.e., *Hampel Filter*, the similar motion morphology-based segmentation, and the PCA-based subcarrier fusion, are used for high-frequency noise elimination, motion segmentation, and data dimension reduction, respectively.

Then, PhaseAnti applies the *Discrete Wavelet Transform* (DWT) on the processed CSI component to obtain the motion shape features. In this part, WAs and IPAs are first distinguished based on the DWT threshold since they have the essential difference that is whether this activity contains position changes. For WAs, due to the fact that different WAs have direct frequency difference of the movement, to further reduce system computation cost, these activities are directly classified by the threshold of the fixed scale DWT coefficients. For IPAs, since the frequency difference between these activities is not apparent and the main difference is reflected on the motion shapes in the time domain, the DWT features are first extracted from these motion shape waveforms and then fed into the classification part for the final activity recognition.

Finally, the classification models are generated by feeding these motion shape features to recognize the specific IPA. PhaseAnti trains an ensemble classifier voted by each classifier of Tx-Rx pair components to generate a classification model for each IPA using the training NLPEV data. We choose the *k-Nearest Neighbor* (kNN) classifier since it essentially searches the entire feature space to match the shape features with high efficiency [11]. Thus, it is most suitable for CSI data classification. Unlike the traditional kNN, PhaseAnti employs the *Dynamic Time Warping* (DTW)-based distance metric while training the kNN classifier to better compare the shape features of any two motions.

4.2 CCI Detection

In non-interfering environments, the channel is only occupied by one device in the range. This device sends and receives frames as usual, and hence the number of received frames per unit time is not decreased and stays almost invariant. Whereas in CCI scenarios, the probability that this device senses an ongoing transmission becomes larger, and the number of delay increases [10], which finally leads to a decrease of the received frame number per unit time. Therefore, by recording the number of received frames per unit time, the CCI existence is judged.

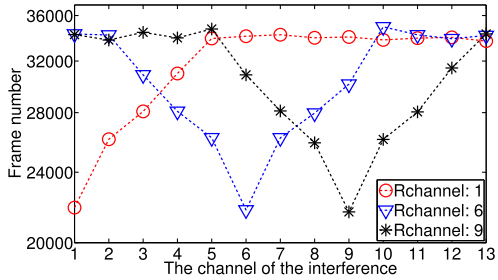


Fig. 5. The number of received frames in six minutes.

Fig. 5 depicts the number of frames received by the recognition laptop in six minutes as the channel of the only one interference AP varies. We repeat each experiment ten times and take the average as the number of received frames per unit time to reduce errors. When the channel of the recognition AP is channel 6, the number of received frames per unit time first decreases rapidly, then rises quickly, and finally stabilizes. This is because CCI exists when the adjacent space between two channels is less than 4, which results in the decreasing of the received frame number per unit time. Similar results are obtained when the recognition AP is set to channel 1 and channel 9. In particular, the CCI existence is determined by the following equation:

$$\frac{|P - S \times t_{re}|}{S \times t_{re}} < TH_{pa}, \tag{9}$$

where P is the number of received frames during time t_{re}, and S represents the sampling rate. TH_{pa} is the threshold corresponding to this equation. Eq. (9) limits that the difference between the number of the received frames during time t_{re} and that received in the non-interfering scenario during time t_{re} should within a given threshold TH_{pa}. TH_{pa} is an empirical value independent of different CCI scenarios. Although the CCI is different in different CCI environments, the non-interference states of different environments are similar. We find that when the TH_{pa} is set to 0.069, any CCI environment that satisfies Eq. (9) does not affect the accuracy of the amplitude component used for motion recognition and can be considered as the non-interfering scenario. Since 0.069 is the dividing line of the CCI intensity whether the amplitude value is affected by CCI in different environments, we set the value of TH_{pa} to 0.069 to determine the existence of CCI.

4.3 Phase Data Processing

Phase Extraction and Unwrapping. As introduced in Section 3, the raw phases at the receiver can be measured according to Eq. (4). However, as shown in Fig. 6, the raw phases distribute between $-\pi$ and π due to the periodicity of the tangent function, which leads to the ambiguity of the subcarrier relationship. Therefore, to recover the real phases of all subcarriers, the raw phases are unwrapped by tracking the corresponding periodic integers of subcarriers [26]. With such unwrapping, the unwrapped phases of each subcarrier become an approximately linear curve.

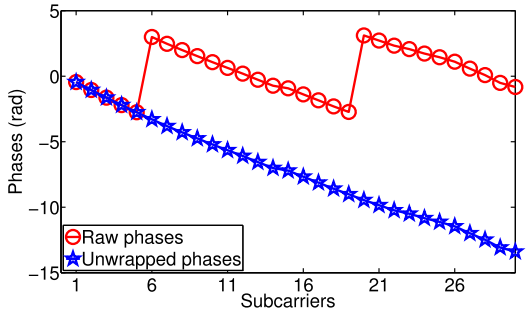


Fig. 6. Phase unwrapping.

Algorithm 1. Phase Filter Based on Gradient Fluctuations

Input: subcarrier phases of frames $\boldsymbol{U} = \{\boldsymbol{\phi}^1, \boldsymbol{\phi}^2, \ldots\}$.
Output: randomly changing phase frame indexes $\boldsymbol{Index}$.

for $i := 1; i <= length(\boldsymbol{U}); i++$ **do**
  $\boldsymbol{G}^i := gradient(\boldsymbol{\phi}^i)$. //Gradient of phases.
  $\boldsymbol{R}^i := gradient(\boldsymbol{G}^i)$.
  $var := 0$. //Variation of gradient.
  for $k := 1; k <= 30; k++$ **do**
    $var := var + abs(R^i_k)$.
  end for
  if $var/(30 \times (\max(\boldsymbol{R}^i) - \min(\boldsymbol{R}^i))) > \hbar_{\max}$ **then**
    $\boldsymbol{Index}.add(i)$.
  end if
end for
return $\boldsymbol{Index}$.

Gradient Estimate and Phase Compensation. Due to the low RSSI of received signals and the unstable CSITOOL [4], not all subcarrier phases are measured accurately and unwrapped successfully [26], [27]. Moreover, these incorrectly measured phases are harmful to HAR due to phase aliasing. Thus, it is necessary to eliminate and then smooth the incorrectly measured phases of the sampled CSI frames. Fig. 7a depicts that some incorrectly unwrapped phases (marked with black triangles) change from one subcarrier to another subcarrier randomly with different gradients. As shown in Fig. 7b, we observe that the gradients of regularly changing phases are more concentrated than the gradients of randomly changing phases (marked with red circles). Also, the gradient variances of randomly changing phases are steep and prominent in all frames (marked with black rectangles). To find the frames with incorrectly measured phases, inspired by the above observation, we design an algorithm based on the phase gradient fluctuation threshold to seek out the frames with larger phase gradient fluctuation. Algorithm 1 describes the searching process. Specifically, the *gradient* is a MATLAB function that calculates the numerical gradient. We first use the *gradient* function twice to calculate the two-step gradient of different subcarrier phases in one frame (e.g., 30 subcarriers for CSITOOL). Then, we calculate the average of the absolute values of the subcarrier two-step gradients. Finally, by employing the *Min-Max Normalization* algorithm to normalize the average, we calculate the final phase fluctuation metric, and its range is between 0 and 1.

From the cumulative distribution of the normalized gradient variances shown in Fig. 7c, we observe that the frames with randomly changing phases (larger variance) are few in all frames. Furthermore, Fig. 7b shows that these frames appear discontinuously. Thus, the frames with randomly changing phases can be smoothed and compensated to capture motions continuously. Specifically, the *moving average filtering* [29] with time weight is exploited to smooth the randomly changing phase frames. The smoothed phases of the frames in the $\boldsymbol{Index}$ can be expressed

$$\boldsymbol{\phi}' = \frac{1}{|Nor|} \sum_{i \in Nor} \boldsymbol{\phi}^i \quad Nor \subseteq W \tag{10}$$

$$|W| = S \cdot t_{act}, \tag{11}$$

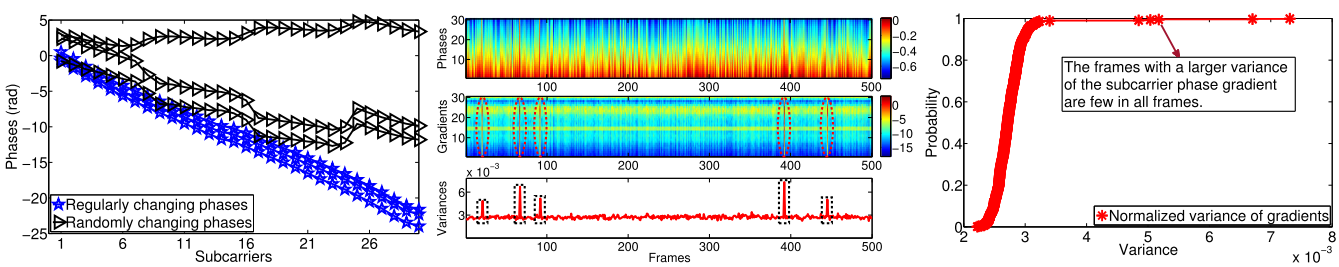
where $\boldsymbol{\phi}'$ are the smoothed subcarrier phases, and W is the window set of the moving average filter, whose size is determined by the sampling rate S and the motion-related time factor t_{act}. To minimize the deviation caused by smoothing, based on empirical knowledge, t_{act} is set to 0.3s. Nor is the set of correctly unwrapped frames in W, and $|Nor|$ is the number of the correctly unwrapped frames in W.

Elimination of Environmental Impacts. As introduced in Section 3.4, the subcarrier phases $\boldsymbol{\phi}$ measured at the receiver can be expressed as

$$\boldsymbol{\phi} = \boldsymbol{\varphi}_{re} + \boldsymbol{L}_{pbd} + \boldsymbol{L}_{sfo} + \boldsymbol{L}_{cfo} + \boldsymbol{L}_{tof} + \boldsymbol{N}_{to}. \tag{12}$$

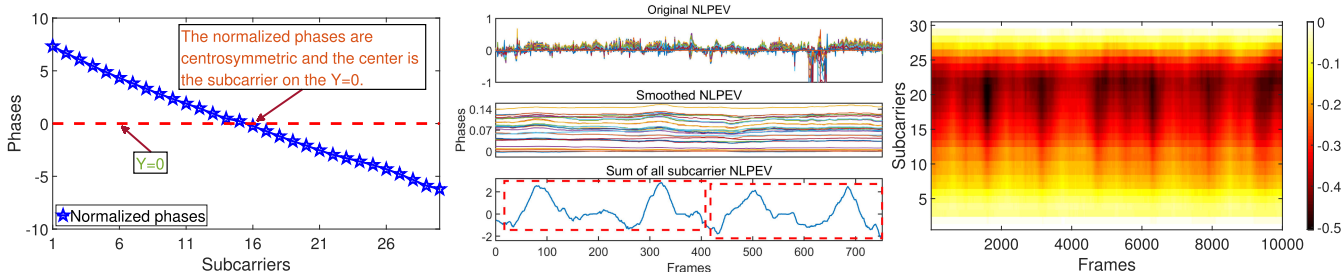
In particular, the phase errors of PBD ($\boldsymbol{L}_{pbd}$) and SFO ($\boldsymbol{L}_{sfo}$) in the same frame are related to the subcarrier index set $\boldsymbol{K}$, which can be represented as

$$\boldsymbol{L}_{pbd} = 2\pi\alpha \cdot \boldsymbol{K} \tag{13}$$



(a) Unwrapped phases of different frames. **(b)** Unwrapped phases, gradients, and normalized gradient variances. **(c)** CDF of normalized gradient variances.

Fig. 7. Unwrapped phase observation and filtering.



(a) Normalized phases after elimination of environmental impacts.

(b) NLPEV for running frames. Subject keeps running during the frame receiving process.

(c) The heat map of NLPEV for running frames.

Fig. 8. Normalized phases and obtained NLPEV data.

$$L_{sfo} = 2\pi\beta \cdot K, \tag{14}$$

where α and β are constant depending on PBD and SFO. The offset of ToF L_{tof} is related to subcarrier frequencies

$$L_{tof} = 2\pi t_f F, \tag{15}$$

where t_f is ToF and affected by device locations. F is the set of subcarrier frequencies represented by the center subcarrier frequency f_c, the identity matrix $\vec{q}$, the frequency difference between two adjacent subcarriers f_b (equals to 312.5 kHz [8]), and the subcarrier index set K, i.e., $F = f_c \cdot \vec{q} + f_b K$. Therefore, the phase error caused by ToF L_{tof} can be rewritten as

$$\begin{aligned} L_{tof} &= 2\pi t_f F \\ &= 2\pi t_f (f_c \cdot \vec{q} + f_b K) \\ &= 2\pi t_f f_c \cdot \vec{q} + 2\pi t_f f_b K \\ &= Z + 2\pi t_f f_b K, \end{aligned} \tag{16}$$

where $2\pi t_f f_c \cdot \vec{q}$ is independent of K. In one frame, t_f is a constant value, thus we use Z to replace the first part. In this respect, the phase measured at the receiver can be reformulated as

$$\begin{aligned} \boldsymbol{\phi} &= \boldsymbol{\varphi}_{re} + \boldsymbol{L}_{pbd} + \boldsymbol{L}_{sfo} + \boldsymbol{L}_{cfo} + \boldsymbol{L}_{tof} + \boldsymbol{N}_{to} \\ &= \boldsymbol{\varphi}_{re} + 2\pi\alpha \cdot \boldsymbol{K} + 2\pi\beta \cdot \boldsymbol{K} + \boldsymbol{L}_{cfo} + \boldsymbol{Z} + 2\pi t_f f_b \boldsymbol{K} + \boldsymbol{N}_{to} \\ &= \boldsymbol{\varphi}_{re} + 2\pi(\alpha + \beta + t_f f_b)\boldsymbol{K} + \boldsymbol{L}_{cfo} + \boldsymbol{Z} + \boldsymbol{N}_{to} \\ &= \boldsymbol{\varphi}_{re} + 2\pi\lambda \cdot \boldsymbol{K} + \boldsymbol{L}_{cfo} + \boldsymbol{Z} + \boldsymbol{N}_{to}. \end{aligned} \tag{17}$$

For a specific frame, λ is a constant value and represents the sum of α, β, and $t_f f_b$. Besides, L_{cfo} is also a constant for each subcarrier in the same frame and can be estimated by $\boldsymbol{\phi}$ [13]. Thus, we use C to replace the sum of L_{cfo} and Z. Then the received phases can be rewritten as

$$\begin{aligned} \boldsymbol{\phi} &= \boldsymbol{\varphi}_{re} + 2\pi\lambda \cdot \boldsymbol{K} + \boldsymbol{C} + \boldsymbol{N}_{to} \\ &= 2\pi\lambda \cdot \boldsymbol{K} + (\boldsymbol{\varphi}_{re} + \boldsymbol{C}) + \boldsymbol{N}_{to} \\ &= 2\pi\lambda \cdot \boldsymbol{K} + \boldsymbol{C}^* + \boldsymbol{N}_{to}, \end{aligned} \tag{18}$$

where C^* contains the real phase φ_{re} and the constant C. According to [13], C^* can be estimated by the phases of a pair of mirror subcarriers measured at the receiver. Specifically, we sum the phases (ϕ_{-1} and ϕ_1) measured at the receiver of a pair of mirror subcarriers -1 and 1 (subcarriers 15 and 16 in CSITOOL) as the following equation:

$$\begin{aligned} \phi_{-1} + \phi_1 &= 2\pi\lambda \cdot (-1+1) + 2 \cdot \boldsymbol{C}^* + N_{to,-1} + N_{to,1} \\ &= 2 \cdot \boldsymbol{C}^* + N_{to,-1} + N_{to,1}, \end{aligned} \tag{19}$$

where $N_{to,-1}$ and $N_{to,1}$ are the NLPEs of subcarrier -1 and subcarrier 1 respectively, and $N_{to,-1} + N_{to,1} \approx 0$ [26]. Therefore, C^* can be calculated approximately as

$$\boldsymbol{C}^* \approx \frac{\phi_{-1} + \phi_1}{2}. \tag{20}$$

Here, C^* is subtracted from the phases of all received frames for the elimination of environmental impacts. In particular, the elimination of environmental impacts is necessary to extract the CCI-independent phase component N_{mo}. This is because different environments cause different LPEs on CSI. If this effect is not filtered out, the remaining part is not only be related to the activity but also related to the specific environment. However, since we want to extract the only activity-related component to perform HAR, it is very important to filter out environmental impacts. As shown in Fig. 8a, after eliminating the environmental impacts, the normalized phases across subcarriers are evenly distributed on both sides of Y = 0 and approximately centrosymmetric.

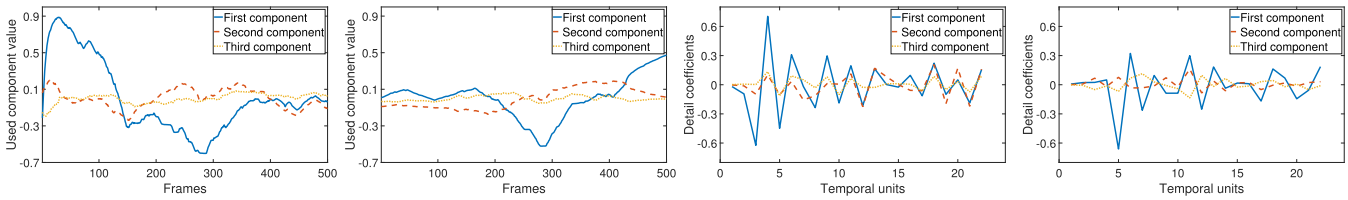
Elimination of Hardware Impacts and Obtain NLPEV. After eliminating the environmental impacts, the total NLPE N_{to} caused by imperfect hardware design and human motion can be expressed as

$$\boldsymbol{N}_{to} \approx \boldsymbol{\phi}_E - 2\pi\lambda \cdot \boldsymbol{K}, \tag{21}$$

where $\boldsymbol{\phi}_E$ denotes the normalized phases after subtracting C^*. Since C^* is obtained through an approximately equal relationship, this processing generates a bias in $\boldsymbol{\phi}_E$. To obtain a relatively steady NLPE N_{to} and to mitigate the bias generated by eliminating the environmental impacts, similar to [26], we use the deviation between the normalized phases $\boldsymbol{\phi}_E$ and the fitted line $\boldsymbol{\gamma}$ to represent the stable NLPE N_{to}^{st}. Particularly, the fitted line $\boldsymbol{\gamma}$ is generated by connecting two points, i.e., $(-28, \phi_{E,-28})$ and $(28, \phi_{E,28})$, which can be expressed as

$$\begin{aligned} \boldsymbol{\gamma} &= slope \cdot \boldsymbol{K} + bias \\ &= \frac{\phi_{E,28} - \phi_{E,-28}}{56} \cdot \boldsymbol{K} + \frac{\phi_{E,-28} + \phi_{E,28}}{2}. \end{aligned} \tag{22}$$

Here $bias$ is not equal to C^* since C^* is the deviation of the curve after unwrapping while $bias$ is the bias of the



(a) Motion waveforms for running. **(b)** Motion waveforms for walking. **(c)** Shape features for running. **(d)** Shape features for walking.

Fig. 9. The motion component waveforms after PCA and the extracted DWT shape features of the selected PCA components, for the walking and running motion samples from the first Tx-Rx pair.

fitted straight line γ. After the fitted line γ is determined, the stable NLPE N_{to}^{st} can be obtained by

$$N_{to}^{st} = \phi_E - \gamma. \tag{23}$$

Since the NLPE caused by the imperfect hardware design N_{ha} is a constant for the specific network card [26] and N_{mo}^{em} equals 0 when there is no motion in the range, we can calculate N_{ha}^{em} in an empty room without human motion

$$\begin{aligned} N_{ha}^{em} + N_{mo}^{em} &= \phi_{E,em} - \gamma_{em} \\ N_{ha}^{em} &= \phi_{E,em} - \gamma_{em}, \end{aligned} \tag{24}$$

where $\phi_{E,em}$ and γ_{em} denote the normalized phases and the fitted line in the empty room, respectively. Then, N_{ha}^{em} is used to represent the NLPE caused by the imperfect hardware design in all frames. Here, we subtract N_{ha}^{em} from the stable NLPE N_{to}^{st} of each received frame for the elimination of hardware impacts. Hence, the NLPEV N_{mo} caused by human movements of each frame can be obtained by

$$\begin{aligned} N_{mo} &= N_{to}^{st} - N_{ha}^{em} \\ &= \phi_E - \gamma - N_{ha}^{em}. \end{aligned} \tag{25}$$

Hampel Filtering and Adaptive Segmentation. Since the DC component and high-frequency noises affect motion segmentation and feature extraction, to obtain robust NLPEV motion samples, we perform the data calibration to remove the DC component and the high-frequency noises. Specifically, we leverage *Hampel Filter* [30] to diminish the high-frequency NLPEV glitches using a sliding window of 500 frames and the threshold of 0.01. Fig. 8b presents the smoothed NLPEV data. The original NLPEV of all subcarriers has high-frequency noises. Nevertheless, after implementing the proposed filtering scheme, the high-frequency noises are removed. Subsequently, we use the dynamic sliding window-based filtering algorithm [24] for the detrend of the NLPEV. After removing the trend of the NLPEV, as shown in Fig. 8b, the sum of the absolute values of each subcarrier NLPEV clearly shows that the same motion (marked as red rectangles) has similar NLPEV fluctuations. Due to the motion information loss, the sum of the absolute values of each subcarrier NLPEV is not suitable for HAR. Thus, Fig. 8b is only a schematic diagram showing that NLPEV periodically fluctuates with the periodic activity. The specific processing of NLPEV is introduced in the *PCA-based Subcarrier Fusion* part and Section 4.5.

Fig. 8c shows the NLPEV series patterns for the running frames, which further verifies that NLPEV has similar waveforms for periodic activities. Therefore, following the Motion-Fi+ [15] work, the goal of our segmentation algorithm is to find optimized start and end points, based on which we can make the difference between each segmentation sample of one activity as small as possible. Since different activity samples have different time durations, we define the similarity of the two activity samples by stretching or compressing these activities to an equal length using DTW. The specific algorithm can be found in Motion-Fi+ [15]. Besides, for the empty room and the activities with small fluctuations, i.e., the standing, and the sitting in the activity set, we use a sliding window with a fixed length of 500 to segment the samples.

PCA-Based Subcarrier Fusion. After the motion segmentation, the CSI motion waveforms of all subcarriers from each Tx-Rx pair are extracted. In particular, we use $V_{i,k}$ to represent the CSI waveform of i^{th} motion sample extracted from the k^{th} Tx-Rx pair. We apply PCA on these CSI waveforms to remove the noisy component and to perform subcarrier fusion for data dimension reduction.

After performing PCA, top ξ principal components for all cases of $V_{i,k}$ are retained, including $\xi - 1$ motion components and one noise component. Considering the tradeoff between the projection contribution level of the component and the computation complexity, we set ξ to 4. Next, we follow the fluctuation variance-based algorithm [11] to determine the location of the noise component. Finally, the noise component is removed accordingly from the top 4 principal components ${V_{i,k}}^{\{1:4\}}$ to get 3 motion component waveforms.

4.4 Amplitude Data Processing

To minimize computational cost, the amplitude information ($|h|$ in Eq. (3)) is employed for realizing HAR in the scenarios without CCI. Similar to NLPEV processing methods, the *Hampel Filter*, the similar motion morphology-based segmentation algorithm, and the PCA-based subcarrier fusion approach, are employed for high-frequency noise filtering, motion segmentation, and data dimensional reduction of the amplitude data $|h|$, respectively.

4.5 Feature Extraction

Due to the fact that the shapes retain both time and frequency domain information of the waveforms and can intuitively reflect the impacts of motions on signals, we thus employ the shapes of the motion waveforms as their features for the activity classification. Figs. 9a and 9b show the motion component waveforms after PCA from the first Tx-Rx pair, and the noisy component is removed accordingly. We can observe that the shape components of different motions are quite distinct from each other after the suitable processing.

In addition, the computational complexity problem needs to be further solved. Particularly, directly using the extracted waveforms as motion features leads to the high computational cost since the component waveforms contain hundreds of data points for one motion sample. Therefore, we apply DWT to compress these waveforms while maintaining most of the time and frequency domain information.

Discrete Wavelet Transform. The DWT of a discrete sampling signal can be written in terms of wavelet basis functions as

$$y[n] = \frac{1}{\sqrt{B}}\sum_{k} \nu(i_0,k)\boldsymbol{\ell}_{i_0,k}(n) + \frac{1}{\sqrt{B}}\sum_{i=i_0}^{\infty}\sum_{k} \varsigma(i,k)\boldsymbol{\rho}_{i,k}(n). \tag{26}$$

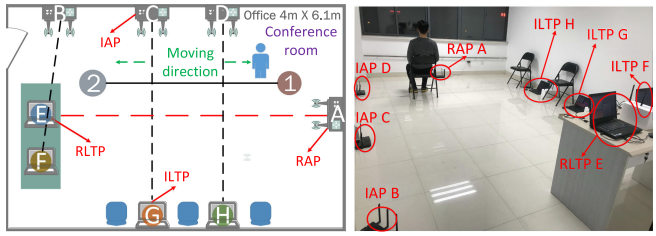
Where B is the length of the signal $y[n]$. The functions $\boldsymbol{\ell}_{i_0,k}(n)$ are scaling functions, and its corresponding coefficients $\nu(i_0,k)$ are known as approximation coefficients. Similarly, the functions $\boldsymbol{\rho}_{i,k}(n)$ are known as wavelet functions, and the corresponding coefficients $\varsigma(i,k)$ are called as detail coefficients. To calculate approximation and detail coefficients, the orthonormal scaling and wavelet functions should be first chosen. Based on the orthonormal basic functions, the approximation and detail coefficients of the i^{th} scale can be expressed as

$$\nu(i,k) = \frac{1}{\sqrt{B}}\sum_{n} y[n]\boldsymbol{\ell}_{i+1,k}(n) \tag{27}$$

$$\varsigma(i,k) = \frac{1}{\sqrt{B}}\sum_{n} y[n]\boldsymbol{\rho}_{i+1,k}(n). \tag{28}$$

In order to perform the satisfactory compression on the selected PCA component waveforms using DWT, the appropriate wavelet and scaling filters are needed. Specifically, we choose *Daubechies D4* wavelet and scaling filters since *Daubechies D4* filters can perform orthogonal transformation faster [11] and have a better ability to characterize the local features of the signal in both time domain and frequency domain.

The DWT structure design is directly determined by the movement frequencies of recognized activities. Considering the motion frequency in the activity set, the average speed of running is about 10km/h, not less than 5km/h, and the corresponding cadence is 170 (2.833Hz) and 85 (1.417Hz), respectively. According to Section 3.2, the true phase at the receiver is a periodic signal with the same frequency as the periodic movement, and thus the CSI frequency response caused by the running motion is ranging from 1.417Hz to 2.833Hz. Similarly, since the walking speed is usually less than 5km/h and its cadence is less than 85 correspondingly, the frequency response of CSI caused by walking is less than 1.417Hz. Besides, for other activities, i.e., sitting and standing, the main motor element of these motions is the heartbeat. Since the heartbeat of people in the stationary state usually ranges from 60 to 110 per minute, the CSI frequency response caused by these motions ranges from 1Hz to 1.833Hz accordingly. Based on the above analysis and combined with the 100 frames/s sampling rate, the DWT structure scales are set to 5. Therefore, the DWT can obtain an approximation coefficient $\nu(5,k)$ $(0 \rightarrow 1.5625Hz)$ and a sequence of detail coefficients $\varsigma(1,k), \varsigma(2,k), \cdots, \varsigma(5,k)$ $(1.5625 \rightarrow 3.125Hz)$.



(a) The schematic diagram. **(b)** The actual scene.

Fig. 10. Evaluation scenario in a conference room.

Activity Type Judgment. The WAs, e.g., walking and running, are difficult to distinguish from each other accurately by traditional methods since the movement steps of these two activities are extremely similar. Nevertheless, these WAs are easy to be distinguished from IPAs since they have the essential difference, which is whether this activity contains position change. Therefore, before classifying the specific activity, the classification between WAs and IPAs is performed. Specifically, the IPAs and the WAs are distinguished by the High-Frequency Energy (HFE) of CSI and can be calculated by the detail coefficients for level 5

$$HFE = \sum_{k=1}^{Len} [\varsigma(5,k)]^2, \tag{29}$$

where *Len* is the length of the detail coefficients $\varsigma(5,k)$. When HFE is larger than the threshold (TH_{st}), the activity is judged as WA. Otherwise, the activity is judged as IPA. Since there are too many factors affecting TH_{st}, even the antenna feeder length and antenna gain can affect this parameter, we cannot quantify the impact of various factors in different environments on TH_{st}. Therefore, to deal with this problem, in a new environment, we collect ten empty room samples and ten walking samples, respectively, to determine the specific threshold in this new environment. Since these samples can be collected in a short time, the threshold value in the new environment can be quickly and easily determined. Specifically, according to Eq. (29), we use the average of the maximum HFE value of ten empty room samples and the minimum HFE value of ten walking samples as the value of TH_{st}. By this way, the value of TH_{st} in the conference room (Fig. 10) is 1.05.

Walking Activity Processing. Since the walking frequency usually less than 1.417Hz and the running frequency normally larger than 1.417Hz, according to the DWT frequency interval of the detail coefficients for level 5 (range from 1.5625 to 3.125Hz), the strength of the CSI frequency response in this interval can directly determine whether it is walking or running. Thus, when motions are judged as WAs, the detail coefficients for level 5 are used to distinguish between walking and running. In particular, the average peak to peak (APTP) value of the detail coefficients for level 5 is employed to recognize the specific WA. When the APTP value is larger than the threshold (TH_{run}), the motion is regarded as running. Otherwise, the motion is classified as walking. Similar to the calculation process of TH_{st}, in a new room, we employ the average value of the maximum

APTP of ten walking samples and the minimum APTP of ten running samples as the value of TH_{run} in this room. According to this calculation method, the TH_{run} in the conference room (Fig. 10) is 0.512. Figs. 9c and 9d show the detail coefficients for level 5 of running and walking samples. Since the movement frequencies of running and walking are different, the fluctuation range and frequency of corresponding coefficients for these motions are obviously different. Running motion has a more significant frequency response on this fixed bandpass filter. Additionally, the motion sequences are significantly shorter after the DWT feature extraction, and the computation complexity is reduced accordingly.

In-Place Activity Processing. When motions are regarded as IPAs, combining the frequency range of these motions, the sum of the approximation and detail coefficients for level 5 is employed to extract the special shape features for each motion. Then, these shape features of IPAs are fed to the classification part for the final activity recognition.

4.6 Classification

After getting the DWT-based shape features of motions, PhaseAnti uses them to establish the training models for classification. Since a comparison metric that provides the valid measurement of similarity between different motion shape features is crucial for classification, PhaseAnti employs DTW to calculate the distance between two waveforms by realizing optimal alignment between them. Then, PhaseAnti trains an ensemble of kNN classifiers using the DWT shape features from all Tx-Rx pairs. Notably, PhaseAnti obtains decisions from all classifiers in the ensemble and employs *majority voting* to get the final selection. Here, the appliance of DTW on the motion shape features and the training processes of the ensemble classifier are explained in detail.

Dynamic Time Warping. DTW is a similarity measurement by obtaining the minimum distance alignment between any two waveforms. Compared with other measurements of similarity, DTW is able to handle waveforms with different lengths and allows a nonlinear mapping of one waveform to another by minimizing the distance between these two waveforms. Compared with *euclidean distance*, DTW can show an intuitive distance between two waveforms by determining the minimum distance warping path between them even if they are shifted or distorted versions of each other. In our experiments, PhaseAnti employs the open-source implementation of DTW in the *Machine Learning Toolbox* (MLT) [31] to realize this algorithm quickly.

Classifier Training. In order to obtain multiple motion characteristics from different Tx-Rx pairs, the separate classifiers for each of theses shape feature components are built. In particular, we build an ensemble of $3 \times M_{Tx} \times M_{Rx}$ classifiers using the kNN algorithm. The number 3 means the three motion shape components projected by PCA from each Tx-Rx pair. M_{Tx} and M_{Rx} are the antenna numbers of Tx and Rx, respectively. Each classifier is trained using the corresponding feature component. To recognize various activities accurately, PhaseAnti puts the motion shape feature components into their corresponding kNN classifiers and obtains a separate decision from each classifier in the ensemble. Each kNN classifier uses the DTW distance metric to search for the majority class label among k nearest neighbors of the corresponding motion shape features. The value of the parameter k is 3. The 10-fold cross-validation algorithm [32] is used in the proposed DTW-based kNN method to avoid over-fitting. Specifically, for each fold in each CCI scenario, nine out of ten samples for one motion are used as templates and one-tenth of the samples for this motion are used for testing. After all folds complete the cycle, each sample is taken as a testing sample once. In order to better compare the impact of different CCI scenarios on CSI and make CCI as the only variable to test system performance, for each CCI scenario in each indoor room, the corresponding templates are all collected from this scenario. Finally, PhaseAnti decides the final choice through *majority voting* on the decisions of all kNN classifiers in the ensemble.

5 Experimental Evaluation

5.1 Implementation

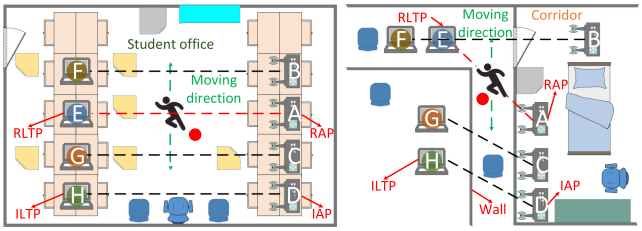
In the experiments, we use two Lenovo laptops as a recognition access point (RAP) and a recognition laptop (RLTP), respectively, both equipped with the Intel 5300 NIC with three antennas. In addition, we employ three TP-link WiFi routers as interference APs (IAPs) and three other laptops as interference laptops (ILTPs). PhaseAnti is implemented on the Ubuntu desktop 14.04 LTS OS for both RAP and RLTP. PhaseAnti exploits the RLTP to collect per frame CSI at a rate of 100 frames per second using CSITOOL [4].

We conduct extensive experiments with 8 persons over one month. As shown in Fig. 10, the $4 \times 6.1m^2$ test room includes a computer table and several chairs. As introduced in Section 4.1, nine daily activities and a reference state, i.e., the empty room, are designed to observe the impacts of CCI on CSI signals with or without human movements. The WAs, i.e., walking and running are performed according to the direction of the green arrow, while the IPAs are performed at point 1.

A comparison experiment is given to measure the impacts of various CCI on CSI, including the following scenarios:

- *Non-interfering scenario:* RAP A is set to channel 1 and forms a wireless link with RLTP E. AP B, laptop F, and other devices are powered off.
- *Simple and constant CCI scenario:* Based on the above setting, AP B is powered on and set to channel 3 since half of channel 3 subcarriers are overlapped with the subcarriers of channel 1. Then, Laptop F is powered on and connected to AP B with a 2.5MB/s CCI interference traffic rate (ITR) (100 frames/s with 25000 bytes packet length).
- *Complex and varying CCI scenario:* In addition to the above settings, laptops G and H are powered on and connected to APs C and D, respectively. The ITRs of laptops G and H are 1MB/s and 5MB/s, respectively. Moreover, IAPs B, C, and D, are adjusted from channel 1 to channel 5 to produce a varying CCI since only channels 1 to 5 overlap with the channel of RAP A.

We examine the performance of single-user activity recognition in three given scenarios. For each participant in



(a) The evaluation scenario in a student office room. **(b)** The evaluation scenario in a corridor.

Fig. 11. Evaluation scenarios in a student office room and in a non-line-of-sight (NLOS) environment with multipath, i.e., a corridor.

each CCI scenario, we collect 54 samples for each activity or state. The activity is performed continuously when we collect the CSI data for this activity. Therefore, a total of $3 \times 8 \times 10 \times 54$ (12960) activity samples are collected for training and testing.

5.2 Experimental Setting of Multipaths and NLOS

To comprehensively evaluate the performance of PhaseAnti in different practical scenarios, we add two scenarios with rich multpath and behind-wall cases, i.e., the student office and the corridor. Specifically, the schematic diagrams of these two scenarios are shown in Figs. 11a and 11b, respectively. The $16.3 \times 10.4m^2$ office room contains multiple desks and chairs. Since the office is big and spacious, we can also test the effective range of one wireless link. Besides, in the corridor, the wireless signal of the RAP passes through the wall and reaches the receiver (RLTP) through multiple paths. Therefore, the corridor scenario can test the recognition performance of PhaseAnti under the electromagnetic condition of complex multipath and NLOS. In particular, as shown in Figs. 11a and 11b, the WAs are executed along the green arrow, while the IPAs are performed at the place marked as the red dot. Three IAPs and three ILTPs are included in these two rooms and the setting of CCI conditions in these rooms is the same as the complex and varying CCI scenario in the conference room (Section 5.1).

Similarly, we evaluate the performance of single-user motion recognition in these two rooms. For every participant of the 8 people in each room, we collect 54 samples of each motion or state. Therefore, a total of $2 \times 8 \times 10 \times 54$ (8640) activity samples are collected for training and testing.

5.3 CCI Detection Evaluation

In this section, we evaluate the performance of the CCI detection algorithm. Specifically, we set up 50 different CCI scenarios in three different rooms, i.e., the conference room, the office room, and the corridor. The 50 scenarios contain 15 non-interfering scenarios and 35 CCI existing scenarios. For each CCI existing scenario in each room, the number of the CCI APs is randomly chosen between 1 and 5 and the ITR of each interference link randomly varies from 2.5MB/s to 10MB/s. For each of the 8 participants in each CCI scenario, we collect 10 samples of each motion or state. Therefore, a total of $50 \times 8 \times 10 \times 10$ (40000) activity samples are collected for training and testing. We employ the PhaseAnti system to process these data, and use 10-fold cross-validation to avoid over-fitting [32].

TABLE 1
The Performance of the CCI Detection Algorithm

CCI type	Non-interfering	CCI exists	Average DAR
DAR	93.33%	100%	98%

Table 1 shows the CCI detection results. The detection accuracy rate (DAR) is calculated by the ratio of the number of the correctly detected CCI scenarios to the number of the total CCI scenarios. We can observe that the proposed CCI detection algorithm in all CCI scenarios has a high average DAR, which can reach 98%. Especially for CCI existing scenarios, the DAR of the algorithm can reach 100%. These results show that through effective statistics and threshold judgment of the number of the received packets per unit time, the existence of CCI in the environment can be effectively detected.

Table 2 depicts the comparison of the recognition performance with or without the CCI detection algorithm. The RAR in this table denotes the average RAR of all 50 different CCI scenarios. In particular, when running the CCI detection algorithm, the system selects the corresponding component (amplitude or NLPEV) for activity recognition according to the CCI detection result. However, when the CCI detection algorithm does not work, the system randomly selects components for activity recognition in different CCI scenarios. As shown in Table 2, with the effective use of the CCI detection algorithm, the average recognition performance of PhaseAnti is significantly improved. This demonstrates that the accurate use of the NLPEV component for HAR in CCI scenarios according to the detection result can significantly improve the recognition performance.

5.4 Motion Segmentation Result

For a given activity, the $|\boldsymbol{\theta}_{re} \cap \boldsymbol{\theta}_{pr}|/\max\{|\boldsymbol{\theta}_{re}|, |\boldsymbol{\theta}_{pr}|\}$ is defined as the segmentation accuracy rate (SAR), where $\boldsymbol{\theta}_{re}$ is the real set of packet indexes for one activity, and $\boldsymbol{\theta}_{pr}$ is the predicted set. $|\cdot|$ represents the length of the set. Accurate segmentation is helpful to the activity classification. Nevertheless, the activity classification accuracy may be larger than the SAR since an activity can be correctly recognized by the classification model even though the start and end of this activity are not precisely detected. If the predicted start and end points are close to the real ones, the majority of the activity data can be extracted and might be adequate for activity classification.

Table 3 shows the SAR of the proposed segmentation method for different activities. The empty room and the activities with small fluctuations, i.e., standing, and sitting, are not in the Table 3, because they are segmented with a fixed sliding window. The segmented samples come from the three different CCI scenarios introduced in Section 5.1. We can observe that all motions in the activity set can be

TABLE 2
The RAR Improvement of the CCI Detection Algorithm

	Without CCI detection	With CCI dection
RAR	92.89%	97.24%

TABLE 3
The Performance of the Motion Segmentation Method

Motion	SAR	Motion	SAR
Push-ups	96.2%	Sit-ups	94.5%
Lift dumbbells	94.3%	Walking	98.2%
Squats	93.5%	Running	97.7%
Stoop-down	93.6%	Total	95.4%

accurately segmented, and the average SAR for all motions can reach 95.4%. The results show that the segmentation algorithm can accurately realize the effective segmentation of the motion CSI series. Compared with IPAs, it is easier to identify the start and end points of the motions with position change, so the WAs (walking and running) have higher SARs.

5.5 NLPEV Robustness Assessment

Invariant to different CCI scenarios, and stable across time are essential attributes for a CCI robustness HAR component. Therefore, we demonstrate the invariance of NLPEV to different CCI scenarios, and time. To avoid the impacts of human movements, all the invariance verification experiments do not contain motions.

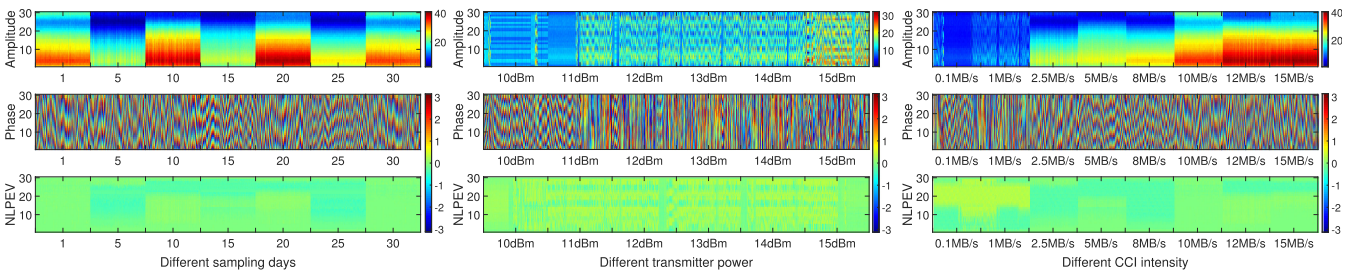
Time Invariance. For HAR systems, the signal component used for recognition should be invariant at different sampling time. However, it is non-negligible that CSIs may change due to different weather conditions, temperatures, and humidity of different days [26]. Since NLPEV is also a phase component extracted from CSI, it is vital to ensure the stability of NLPEV across time. Fig. 12a shows amplitude, phase, and NLPEV at different time without human motion. We measure the CSIs in the conference room, as shown in Fig. 10, and extract NLPEVs day and night for a month. We can see that the amplitude changes with different sampling days and the phase changes randomly due to phase errors. However, NLPEVs are relatively stable and invariant across time.

Transmitter Power Invariance. In order to examine the invariance of NLPEV to the change of transmitter power caused by APs competing for the channel, we change the NIC power of RLTP E from 10dBm to 15dBm (maximum power of 5300 NIC) and fix the power of RAP A. Then, the sampled CSIs and the extracted NLPEVs in different transmitter power are shown in Fig. 12b. It is evident that the amplitude increases with the rising transmitter power. However, the NLPEV changes negligibly with the increasing transmitter power, which clearly demonstrates that the NLPEV is more stable and independent of the varying signal strength.

CCI Intensity Invariance. The increase of CCI intensity causes a decrease in the sampling rate and the confusion of the receiving packet order [10]. Thus, the stability of CSI is affected. To evaluate the impacts of this factor on NLPEV, we keep one IAP (AP B in Fig. 10) and adjust ping rates and packet lengths of the interferer to generate different ITRs. Fig. 12c depicts that the amplitude increases significantly due to the increasing transmitter power caused by APs competing for the channel in more severe CCI scenarios. Nevertheless, the NLPEV stays almost constant even if the CCI ITR is changing. This experiment confirms that NLPEV is invariant to various CCI intensities.

Invariance Performance Evaluation. To quantitatively evaluate the invariance of NLPEV to environmental factors, we also compare the RARs of different signal components under different sampling time, different transmitter power, and different CCI intensities. Take different sampling time as an example, for each participant of the 8 people in each sampling day (days 1, 5, 10, 15, 20, 25, 30 as shown in Fig. 12a), we collect 54 samples of each motion or state. Therefore, a total of $7 \times 8 \times 10 \times 54$ (30240) activity samples are collected for training and testing. We employ the PhaseAnti system to process these data, and use 10-fold cross-validation to avoid over-fitting [32]. The experiment method of different transmitter power and different CCI intensities is same with that of the different sampling days.

Fig. 13 shows the RARs of different CSI components under different influencing factors. The RAR can directly reflect the robustness of the recognition component to these influencing factors. The smaller reduction in the RAR under one environmental factor impact indicates that the component is more robust to this environmental factor. As shown in Fig. 13, since a lot of phase errors are contained, the original phase cannot accurately recognize the motion under any circumstances. Besides, the amplitude component is relatively robust to different sampling times, and achieve a satisfactory RAR in different sampling days. However, under the CCI impacts, i.e., different transmitter power, and different CCI intensities, the RAR of the amplitude component is obviously degraded. In contrast, since the components affected by the CCI are filtered out and the only motion-related HAR component is extracted, PhaseAnti can not only maintain a high RAR stably in different sampling times, but also recognize motions accurately in different



(a) NLPEV remains invariant across the time. **(b)** NLPEV remains invariant under the different transmitter power. **(c)** NLPEV remains invariant under the different ITRs.

Fig. 12. Observation of the properties of NLPEV and other components in various CCI scenarios. Each CCI scenario contains 500 consecutive CSI frames, and the sampling rate is 100Hz.

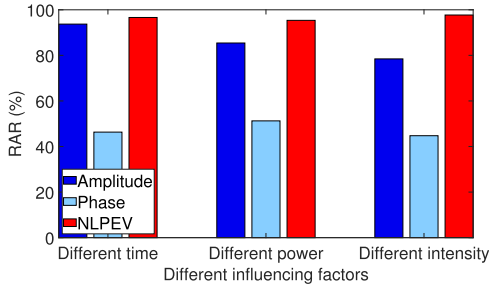


Fig. 13. The performance of different components under different influencing factors.

CCI environments. This shows that compared with the original phase and the amplitude, the proposed component NLPEV has better invariance to the environment change caused by CCI.

These observations demonstrate that the amplitude-based systems can hardly realize anti-CCI due to the unstable subcarrier amplitudes caused by varying CCI, which seriously downgrade the system classification performance. The original phase does not contain enough motion information due to lots of contained phase errors. Nevertheless, the proposed CSI phase component NLPEV remains invariant across various CCI and time. Moreover, the NLPEVs caused by different human activities are distinct (shown in Fig. 9). Therefore, by extracting NLPEV data from the CSI phase to perform HAR, PhaseAnti can achieve anti-CCI.

5.6 Recognition Performance

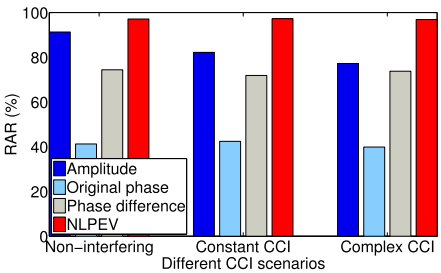
The recognition performance measurement RAR used in this paper is defined as the ratio of the correctly classified motion number Num_{cor} to the whole testing motion number Num_{whole}, which can be expressed as follows:

$$RAR = \frac{Num_{cor}}{Num_{whole}} \times 100\%. \tag{30}$$

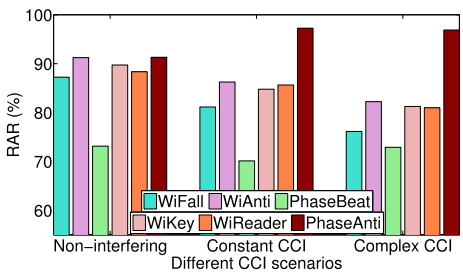
Performance of Different Components. We first compare the recognition performance of different CSI components, i.e., amplitude, original phase, and phase difference data [24], with the proposed component NLPEV in various CCI scenarios. To perform a fair comparison, except for respective signal component extraction processes, all the processing, feature extraction, and classification methods for baseline components are the same as that of the NLPEV component. Furthermore, the training and testing motion samples for different CSI components are the same.

Fig. 14a depicts the RARs of different CSI components in various CCI scenarios. Although the processing methods for all components are the same, we can observe that NLPEV still achieves better performance compared with other CSI components in all CCI scenarios. Since the CCI environment for each component is consistent, higher RARs mean less impact of CCI on the proposed component and better ability to maintain motion information. The original phase has low RARs since many irrelevant phase errors are contained. The performance of the phase difference is not good because not all error components in the phase difference data are effectively filtered out. In addition, the amplitude component achieves a high RAR in the non-interfering scenario, reaching 91.31%. However, the performance of the amplitude component has a significant decrease in CCI scenarios. Furthermore, the more complex CCI environment further downgrades the amplitude component performance. This obviously shows that the amplitude information is indeed affected by CCI since each AP changes the transmission power to better compete for the channel in the CCI environment, which makes the amplitude difference between different samples of the same motion more significant. Nevertheless, the RAR of the NLPEV component does not decrease in CCI scenarios and it reaches 96.89% even in the complex CCI scenario. Moreover, in the non-interfering scenario, the NLPEV component also achieves a higher RAR than the amplitude information. This demonstrates that the NLPEV contains more motion information than the amplitude, even without the impact of CCI.

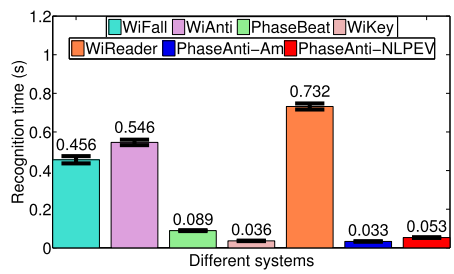
Performance of Different HAR Systems. To verify the performance of different systems in CCI scenarios, we compare the proposed system with the pioneer amplitude-based fall detection system WiFall [5], the state-of-the-art amplitude-based keystroke recognition system WiKey [11], the novel handwriting recognition system WiReader [33], the state-of-the-art anti-CCI system WiAnti [9], and the pioneer phase-based system PhaseBeat [24]. All these baseline systems are based on commercial WiFi devices without adding any additional devices, and these recognition systems can be built through a single AP and a single client, which are same with the implementation of PhaseAnti. For a fair comparison, all these WiFi-based HAR systems use the same training data to train their corresponding recognition models and the same testing data to measure their recognition performance. In particular, these baseline systems use their own CSI processing methods, i.e., different preprocessing methods, different feature extraction methods,



(a) Comparison of different CSI component RARs in various CCI scenarios.



(b) Comparison of different system RARs in various CCI scenarios.



(c) Comparison of different system recognition time for one motion sample.

Fig. 14. Comparison of different HAR systems in the conference room. PhaseAnti uses the amplitude information to perform HAR in non-interfering scenarios.

and different classifiers, to recognize the motions in the activity set.

Fig. 14b reports the performance of different WiFi-based HAR systems in various CCI scenarios. Compared with WiFall, the RARs of WiKey and WiReader are significantly improved in the non-interfering scenario due to the use of advanced signal processing methods and classification technologies. However, in the constant CCI scenario and the complex CCI scenario, CCI significantly downgrades the recognition performance of WiKey and WiReader, which indicates that these novel WiFi-based HAR systems are not robust to CCI. Although only three amplitude components are used in the non-interfering scenario, PhaseAnti achieves a similar RAR with WiAnti using 6 subcarriers and outperforms WiFall using six subcarriers. This demonstrates that the amplitude component used in PhaseAnti can accurately recognize the motion in the non-interfering scenario, and reduce the extraction time of the recognition component as much as possible, which finally improves the recognition efficiency. Furthermore, in CCI scenarios, PhaseAnti achieves more satisfactory RARs and outperforms all the baselines by adopting the NLPEV component to realize HAR. This shows that the recognition ability of the PhaseAnti system in the CCI environment is improved significantly by extracting the CCI-independent component for behavior recognition.

Recognition Time. The recognition time for each motion sample of all systems is shown in Fig. 14c. The recognition time for each system contains the time of signal processing, feature extraction, and classification. Compared with WiFall using six subcarriers, WiAnti has a slower recognition speed because it takes a lot of time to calculate the subcarrier correlation. Nevertheless, under the same processed data volume, PhaseAnti-NLPEV can realize a $10.3\times$ faster speed than WiAnti in CCI scenarios. Moreover, in non-interfering scenarios, the recognition speed of PhaseAnti (PhaseAnti-Am) is further boosted when using the amplitude data to perform HAR. Compared with all baseline methods, PhaseAnti-Am has the fastest recognition speed and the average recognition time of PhaseAnti-Am for one motion sample is 0.033s. Instead of increasing much calculation cost to achieve the CCI robustness, the PhaseAnti system use more time-efficient way to accurately recognize the activities in CCI scenarios. Specifically, by employing the CCI-independent component to perform HAR and effective data dimensionality reduction, PhaseAnti can accurately recognize the motion in the CCI environment, reduce the computational complexity as much as possible, and further achieve a faster recognition speed.

Performance in Complex Multipath Scenarios. We add two rooms with rich multipath and behind-wall cases, i.e., the student office and the corridor, to comprehensively evaluate the PhaseAnti performance in complex electromagnetic environment. The experimental settings are described in Section 5.2.

Fig. 15 shows the performance of the PhaseAnti system and other baseline methods in various multipath indoor environments. The performance of PhaseAnti in the office room is similar to that in the conference room. The results show that the large-scale indoor range can not affect the effective recognition of PhaseAnti. Compared with the environments with less propagation paths, i.e., the conference room and the office room, the rich multipath and the wall in the corridor downgrade the recognition performance of each system. Nevertheless, since PhaseAnti is not affected by the varying CCI in electromagnetic environment, PhaseAnti can still recognize daily motions accurately in the multipath environment (corridor) with a high RAR, reaching 95.67%, and consistently realize the best performance in all multipath scenarios. Furthermore, the multipath effect does not significantly degrade the recognition performance of PhaseAnti, which is reflected in the small reduction of RAR. Since more motion information is maintained due to the valid error component elimination, PhaseAnti is more robust to various rich multipath environments.

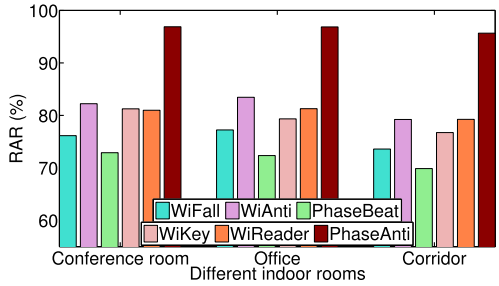


Fig. 15. The performance of different recognition systems in various indoor rooms.

Performance of Processing Algorithms. In order to verify the recognition performance improvement brought by the proposed DWT feature extraction algorithm and the ensemble DTW-based classifier, we design a set of comparison experiments including the experiments with or without the proposed DWT feature extraction algorithm and the experiments with or without the ensemble DTW-based classifier. When the proposed DWT feature extraction algorithm is not used, the system extracts 10 common features, which are widely used in existing WiFi-based HAR systems. The specific features are mean, variance, maximum, minimum, median, first quartile, third quartile, information entropy, spectrum energy, and maximum frequency domain. Besides, when the ensemble DTW-based classifier is not used, PhaseAnti employs the general *Support Vector Machine* (SVM) classifier to realize quick classification. Similar to Section 5.1, all motion data in this part are collected from the conference room, and the corresponding CCI scenario is the complex and varying CCI.

Table 4 shows the improvement of the proposed feature extraction algorithm and the ensemble DTW-based classifier for the PhaseAnti system. Compared with the traditional feature extraction, the PhaseAnti performance (without DTW) is significantly boosted by employing the pertinent DWT feature extraction algorithm and the RAR is increased from 92.11% to 95.35%. This demonstrates that the targeted analysis of the motion frequency can significantly enhance the system recognition performance. Similarly, the PhaseAnti performance (without DWT) is also improved by using the ensemble DTW-based classification method, which shows that the multi-angle voting can capture more motion characteristics to improve the recognition performance. As for the recognition speed, the use of the pertinent DWT feature extraction algorithm significantly boosts the recognition speed of the PhaseAnti system (without DTW)

TABLE 4
The Performance of the Proposed Feature Extraction Algorithm and the Ensemble DTW-Based Classifier in the Complex and Varying CCI Scenario

Measurement	Without DWT	Without DTW	Without DWT and without DTW	PhaseAnti
RAR	93.17%	95.35%	92.11%	96.89%
Recognition time	0.196s	0.064s	0.218s	0.053s

and the recognition time for one motion sample is decreased from to $0.218s$ to $0.064s$. This clearly shows that the valid data dimensional reduction in the wavelet decomposition decreases the computational complexity and finally improves the recognition speed. Besides, the ensemble DTW-based classifier is also helpful to improve the recognition speed of the PhaseAnti system. Therefore, the proposal of the pertinent DWT feature extraction algorithm and the ensemble DTW-based classifier is beneficial for both the recognition accuracy and the recognition speed.

6 Conclusion

In this paper, we have presented PhaseAnti, an anti-CCI HAR system based on WiFi CSI. The core of PhaseAnti is the use of the HAR component NLPEV. In particular, under CCI scenarios, by eliminating irrelevant errors, the CCI-independent phase component NLPEV has been leveraged from the COTS WiFi device since this component remains constant for different CCI scenarios and contains motion information. We have also introduced a suitable calibration method to ensure that NLPEV data are stable and sensitive to the motions. Then, the PCA-based subcarrier fusion algorithm and the DWT-based feature extraction method are employed with minor information loss in order to reduce the data dimension. By using these extracted NLPEV shape features to realize HAR in CCI scenarios, the activity can be recognized accurately with low recognition time. Besides, in non-interfering scenarios, the amplitude information is employed to further reduce the system computation cost. Extensive experiments with different motions have been implemented in various CCI scenarios, and the results show that PhaseAnti can achieve superior performance on RAR and recognition time over existing systems in all cases.

References

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, Third Quarter 2013.

[2] W. Zhu, J. Hu, G. Sun, X. Cao, and Y. Qiao, "A key volume mining deep framework for action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1991–1999.

[3] G. Laput and C. Harrison, "Sensing fine-grained hand activity with smartwatches," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–13.

[4] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, 2011.

[5] Y. Wang, K. Wu, and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 581–594, Feb. 2017.

[6] Y. Gu, F. Ren, and J. Li, "PAWS: Passive human activity recognition based on WiFi ambient signals," *IEEE Internet of Things J.*, vol. 3, no. 5, pp. 796–805, Oct. 2016.

[7] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun.*, 2015, pp. 1472–1480.

[8] *IEEE Standard for Information Technology, IEEE Standard 802.11n-2009*, pp. 1–565, Oct. 2009.

[9] J. Huang *et al.*, "Towards anti-interference human activity recognition based on WiFi subcarrier correlation selection," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6739–6754, Jun. 2020.

[10] J. Huang, B. Liu, H. Jin, and Z. Liu, "WiAnti: An anti-interference activity recognition system based on WiFi CSI," in *Proc. IEEE Int. Conf. Internet of Things IEEE Green Comput. Commun. IEEE Cyber Phys. Soc. Comput. IEEE Smart Data*, 2018, pp. 58–65.

[11] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Recognizing keystrokes using WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1175–1190, May 2017.

[12] Z. Shi, J. A. Zhang, Y. D. R. Xu, and Q. Cheng, "Environment-robust device-free human activity recognition with channel-state-information enhancement and one-shot learning," *IEEE Trans. Mobile Comput.*, pp. 1–1, 2020, doi: 10.1109/TMC.2020.3012433.

[13] N. Yu, W. Wang, A. X. Liu, and L. Kong, "QGesture: Quantifying gesture distance and direction with WiFi signals," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 1, 2018, Art. no. 51.

[14] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38.

[15] N. Xiao, P. Yang, Y. Yan, H. Zhou, X.-Y. Li, and H. Du, "Motion-Fi +: Recognizing and counting repetitive motions with wireless backscattering," *IEEE Trans. Mobile Comput.*, vol. 20, no. 5, pp. 1862–1876, May 2021.

[16] X. Fan, W. Gong, and J. Liu, "TagFree activity identification with RFIDs," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 1, 2018, Art. no. 7.

[17] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.

[18] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.

[19] D. Wu *et al.*, "FingerDraw: Sub-wavelength level finger motion tracking with WiFi signals," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 4, no. 1, pp. 1–27, 2020.

[20] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi: Sign language recognition using WiFi," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 1, 2018, Art. no. 23.

[21] Y. Zheng *et al.*, "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proc. 17th Annu. Int. Conf. Mobile Syst. Appl. Services*, 2019, pp. 313–325.

[22] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–25, 2018.

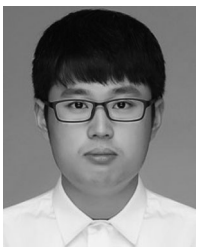
[23] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring motion direction using commodity Wi-Fi for interactive exergames," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 1961–1972.

[24] X. Wang, C. Yang, and S. Mao, "PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices," in *Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst.*, 2017, pp. 1230–1239.

[25] Y. Zeng, D. Wu, R. Gao, T. Gu, and D. Zhang, "FullBreathe: Full human respiration detection exploiting complementarity of CSI phase and amplitude of WiFi signals," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 3, 2018, Art. no. 148.

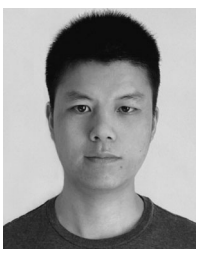
[26] P. Liu, P. Yang, W.-Z. Song, Y. Yan, and X.-Y. Li, "Real-time identification of rogue WiFi connections using environment-independent physical features," in *Proc. IEEE Conf. Comput. Commun.*, 2019, pp. 190–198.

[27] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving accurate CSI phases with commodity WiFi devices," in *Proc. IEEE Conf. Comput. Commun.*, 2017, pp. 1–9.
[28] J. Hua, H. Sun, Z. Shen, Z. Qian, and S. Zhong, "Accurate and efficient wireless device fingerprinting using channel state information," in *Proc. IEEE Conf. Comput. Commun.*, 2018, pp. 1700–1708.
[29] J. S. Hunter, "The exponentially weighted moving average," *J. Qual. Technol.*, vol. 18, no. 4, pp. 203–210, 1986.
[30] R. K. Pearson, "Outliers in process modeling and identification," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 1, pp. 55–63, Jan. 2002.
[31] J.-S. R. Jang, "Machine learning toolbox," 2014. [Online]. Available: mirlab.org/jang/matlab/toolbox/machineLearning, (Dec 1, 2014)
[32] Y. Liu and S. Liao, "Preventing over-fitting of cross-validation with kernel stability," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2014, pp. 290–305.
[33] Z. Guo, F. Xiao, B. Sheng, H. Fei, and S. Yu, "WiReader: Adaptive air handwriting recognition based on commercial WiFi signal," *IEEE Internet of Things J.*, vol. 7, no. 10, pp. 10 483–10 494, Oct. 2020.



Jinyang Huang received the BEng degree from Anhui University, Hefei, China, in 2017, and he is currently working toward the PhD degree in the School of Cyberspace Security, University of Science and Technology of China, Hefei, China. His research interests include human-computer interaction, wireless sensing, wireless communication, and machine learning.

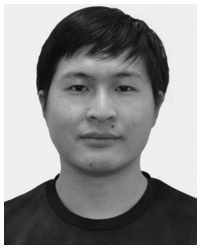
Bin Liu received the BS and MS degrees in electrical engineering from the University of Science and Technology of China, Hefei, Anhui, China, in 1998 and 2001, respectively, and the PhD degree in electrical engineering from Syracuse University, Syracuse, New York, in 2006. Currently, he is an associate professor with the School of Information Science and Technology, University of Science and Technology of China. His research interests include signal processing and communications in wireless sensor and body area networks.



Chenglin Miao (Member, IEEE) received the PhD degree in computer science and engineering from the State University of New York at Buffalo, Buffalo, New York, in 2020. He is currently an assistant professor with the Department of Computer Science, University of Georgia. His research interests include security and privacy, Internet of Things, and machine learning. He is a member of the ACM.

Yan Lu received the bachelor's degree in telecommunication engineering from Anhui Polytechnic University, Wuhu, China, in 2017, and the master's degree in electronics and communications engineering from the University of Science and Technology of China, Hefei, China, in 2020. Currently, he is working toward the PhD degree at the University of Sydney, Sydney, Australia, and his research interests include computer vision and deep learning.

Qijia Zheng received the BE degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2020. He is currently working toward the graduate degree in the School of Cyberspace Security, University of Science and Technology of China, Hefei, China. His research interests include wireless sensing, wireless communication, and machine learning.



Yu Wu received the BS and ME degrees in information security and electrical engineering from the University of Science and Technology of China, Hefei, China. He is currently working toward the PhD degree at Rutgers University, New Brunswick, New Jersey. His research interests include human activity recognition, human-computer interaction, and machine learning theory.

Jiancun Liu received the BS degree from the North China Electric Power University, Beijing, China, in 2017. He is currently working toward the PhD degree in the School of Data Science, University of Science and Technology of China (USTC), Hefei, China. His research interests include software-defined networks, network function virtualization, edge computing, and federated learning.

Lu Su (Member, IEEE) received the MS degree in statistics and the PhD degree in computer science, both from the University of Illinois at Urbana-Champaign, Champaign, Illinois, in 2012 and 2013, respectively. He is currently an associate professor with the School of Electrical and Computer Engineering, Purdue University. His research interests include Internet of Things and cyber-physical systems, with a current focus on wireless, mobile, and crowd sensing systems. He has also worked with IBM T. J. Watson Research Center and National Center for Supercomputing Applications. He has published more than 100 papers in referred journals and conferences, and serves as an associate editor of the *ACM Transactions on Sensor Networks*. He is the recipient of NSF CAREER Award, University at Buffalo Young Investigator Award, ICCPS'17 Best Paper Award, and the ICDCS'17 Best Student Paper Award. He is a member of the ACM.

Chang Wen Chen (Fellow, IEEE) is currently a professor of computer science and engineering with the State University of New York at Buffalo, Buffalo, New York. Previously, he was Allen S. Henry endowed chair professor with the Florida Institute of Technology from 2003 to 2007, a faculty member with the University of Missouri-Columbia from 1996 to 2003 and with the University of Rochester, Rochester, New York, from 1992 to 1996. He has been the editor-in-chief of the *IEEE Transactions on Multimedia* since 2014. He has also served as an editor of the *Proceedings of IEEE*, *IEEE Transactions on Multimedia*, *IEEE Journal on Selected Areas in Communications*, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, and *IEEE Multimedia Magazine*. He is an SPIE fellow. Since 2017, he serves as dean of the School of Science and Engineering, Hong Kong Polytechnic University, Shenzhen.