IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 October 2022, accepted 13 November 2022, date of publication 18 November 2022,
date of current version 23 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3223437

RESEARCH ARTICLE

Human Activity Recognition With Commercial WiFi Signals

CHEN TIAN[1], (Student Member, IEEE), YUE TIAN[1], (Member, IEEE), XIANLING WANG[1], (Member, IEEE), YAU HEE KHO[2], (Senior Member, IEEE), ZHENZHE ZHONG[4], (Member, IEEE), WENDA LI[3], (Member, IEEE), AND BAIYUN XIAO[1], (Student Member, IEEE)

[1]School of Opto-Electronic and Communication Engineering, Xiamen University of Technology, Xiamen 361024, China
[2]School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6012, New Zealand
[3]School of Science and Engineering, University of Dundee, DD1 4HN Dundee, U.K.
[4]Fujian Key Laboratory of Industrial Internet and IoT, Xiamen Intretech Inc., Xiamen 361006, China

Corresponding authors: Yue Tian (yue.tian.xmut@outlook.com) and Xianling Wang (xianling.wang@ieee.org)

This work was supported in part by the Youth Innovation Foundation of Xiamen under Grant 3502Z20206067; in part by the Natural Science Foundation of Xiamen under Grant 2022FCX012503010125; in part by the Natural Science Foundation of China under Grant 62201482; in part by the Natural Science Foundation of Fujian Province, China, under Grant 2019J01874, Grant 2021J011219, Grant 2022J011276, and Grant 2023J01130536; in part by the Innovation Foundation of Xiamen Sage Tech. under Grant ZK-HX21124; and in part by the Xiamen Key Science and Technology Program under Grant 3502Z20221026.

ABSTRACT The next generation of mobile communication aims to extend the capabilities of traditional communication by reshaping the environment with wireless signals. The channel state information can describe the propagation characteristics in wireless communications, which is beneficial in developing wireless communication networks towards intelligent communication and wireless sensing networks. Raspberry PI with Nexmon firmware patched can extract the channel state information from WiFi signals and realize human activity recognition. However, the phase values on some carriers are susceptible to noise, resulting in phase errors after singular value decomposition. To solve this problem, a method is proposed in this paper to find the optimal phase value by dynamic time warping algorithm utilizing the property of orthogonality between amplitude and phase. In contrast to the conventional recognition strategies, the proposed optimal phase extraction method with commercial WiFi signals can further improve the accuracy of the recognition strategy under different complicated scenarios.

INDEX TERMS Wireless sensing, channel state information, WIFI, dynamic time warping.

I. INTRODUCTION

The proposal of the metaverse promotes the development of human-computer interaction technology. Human activity recognition plays a crucial role in it. Nowadays, the widely used human activity recognition is based on AI vision-based and dedicated sensors. Although the above methods can accurately and efficiently recognize the activity, they all have unique application scenarios. The method based on Al vision has high requirements on the brightness of the surrounding environment and can only be used within the Line of Sight (LoS), which may lead to the disclosure of information. The method based on dedicated sensors requires expensive equipment, and the device must be worn on the body, causing great inconvenience in activities. In recent years, WiFi-based human activity recognition has overcome these limitations. The WiFi-based human activity recognition leverages integrated sensing and communication (ISAC) capabilities to reshape human activity. This WiFi-based contactless wireless sensing system has the following advantages: firstly, a large number of ubiquitous WiFi signals can be used for sensing and not be limited by LoS conditions; secondly, people do not need to wear any sensing devices; lastly, the whole system is easy to establish and low cost.

Initially, the WiFi-based activity recognition was based on the Received Signal Strength Indicator (RSSI) [1], [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Ding Xu.

However, RSSI is sensitive to the environment, which makes it unable to provide accurate classification results. By comparing the performance of RSSI and Channel State Information (CSI), Yang et al. [3] proved that CSI could distinguish the multipath effect. Therefore, most existing WiFi-based human recognition systems use CSI for sensing. For instance, the amplitude of CSI can be used to recognize human activities [4], [5]. Zhang et al. [6] proposed the concept of a Fresnel Zone in the wireless sensor field. They showed that the same activity at different locations might lead to significant changes in characteristics. Therefore, it is not accurate enough to judge only from the amplitude of CSI. Zhang et al. [7] summarized the current development status of CSI, including the existing practical and theoretical problems, and provided solutions for future development. All these research works offered significant help for the development of WiFi-based contactless wireless sensing.

Nonetheless, the extraction of CSI was mainly realized using the computer equipped with Intel 5300 network card, which can only collect CSI of 30 subcarriers and cannot meet the requirements of fine-grained recognition.

The WiFi-based activity recognition established by Raspberry PI with Nexmon patched [8], [9], [10] can address the above problems. The system supports IEEE 802.11ac and allows 80 MHz of bandwidth to extract CSI of 256 subcarriers, more than the sum of the subcarriers collected by WiFi chip on the Intel 5300 network card.

However, in the process of collecting CSI, the phase of CSI was easily disturbed by noise, which makes the collected phase value useless. Therefore, this paper put forward an optimal phase method to solve the above problem. The main contributions of this paper are summarized as follows:

1) A contactless wireless system was built on Raspberry PI and completed the recognition of standing, bending over, squatting and holding up hands, which are four common activities in the metaverse.
2) In the process of CSI collection, the phase on some carriers will seriously interfere, resulting in the deviation of the collected CSI phase value. Therefore, this paper proposed a method to find the optimal phase value by dynamic time warping (DTW) algorithm utilizing the property of orthogonality between amplitude and phase.
3) Through testing in different scenarios, the evaluation results proved that the method proposed in this paper performed better than using only the amplitude or using all the phase values.

II. RELATED WORK

At present, many scholars have made research on wireless sensing. WiFall [13] used CSI to recognize human falling. WiFinger [14] devised an environmental noise removal mechanism by filtering out the environmental noise while keeping the CSI pattern resulting from the finger gesture and recognizing the gesture. FarSense [15] employed the ratio of CSI reading from two antennas whose noise was mainly canceled out by the division operation to increase the sensing range significantly. WiDance [16] used antenna diversity to remove random phases and then correlated doppler shift with motion direction to recognize. WiDar2.0 [17] was the first passive tracking system requiring only one WiFi link. It took the angle of arrival, time of flight, and doppler shift into account and achieved submeter level tracking accuracy. WiGesture [18] shifted observation from the traditional transceiver to the hand-oriented view and extracted a novel position-independent feature named Motion Navigation Primitive (MNP). MNP exploited the moving direction of hands, which made gestures easily identified. WiDar3.0 [19] used velocity profiles of gestures from doppler frequency shift (DFS) on at least three links, which acted as unique indicators of gestures for recognition. WiFi-based wireless sensing systems require plenty of experiments to improve accuracy. GaitID [31] extracted specific features of the target so that recognition does not require any restriction on walking trajectory and speed. On this basis. GaitSense [32] extracted target features to better represent gait, and used transfer learning and data augmentation techniques to effectively reduce the requirements on training samples.

Most of the above research works were based on CSI on some carriers, but not all. This paper proposed a system to collect CSI on all the carriers. However, the phase of some carriers may be seriously disturbed which affects the accuracy of recognition. To alleviate this problem, a method finding the optimal phase according to the orthogonal relationship between amplitude and phase has been proposed. The optimal phase and amplitude values were used for activity recognition. The estimation results showed that the proposed method could improve the accuracy of the recognition strategy with commercial WiFi signals under different complicated scenarios.

III. PRINCIPLE

This part mainly describes the basic principles of wireless sensing, including the wireless sensing model, Fresnel Zone, and the extraction of optimal phase matching strategy.

A. WIRELESS SENSING MODEL

Typically, a transmitter sends information to a receiver through a wireless channel. CSI describes the wireless channel by giving each subcarrier's channel frequency response (CFR). The wireless communication model can be characterized by:

$$y(f,t) = H(f,t)\,x(f,t) + n, \tag{1}$$

where $y(f,t)$ is the information received by the receiver; $H(f,t)$ is the transmission matrix of the channel; $x(f,t)$ is the information sent by the transmitter; n is the noise in the channel.

When people move, for example when doing certain activities, the wireless signals in the environment will be reflected,

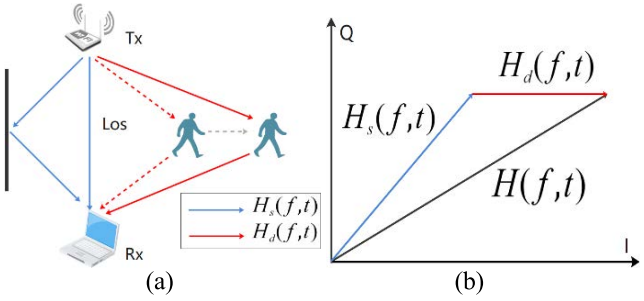


FIGURE 1. The composition of $H(f,t)$. (a) static component $H_s(f,t)$ includes a constant reflection path component and LoS component as indicated by the blue line, while dynamic component $H_d(f,t)$ represents the reflection path changes as indicated by the red line; (b) the phasor representation of $H(f,t)$ as a result of a composition of the static component $H_s(f,t)$ and dynamic component $H_d(f,t)$.

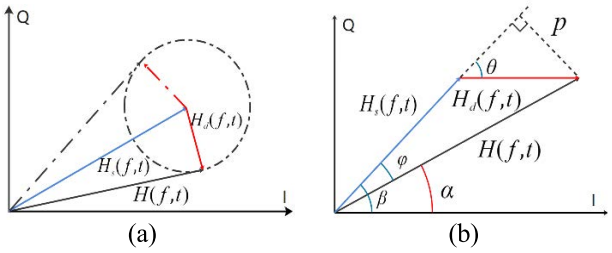


FIGURE 2. Derivation of the formula of the CSI phase when $|H_d(f,t)|$ is a constant: (a) shows the one possibility for $H(f,t)$. (b) shows the parameter description of formula derivation.

diffracted and scattered, resulting in a change in the dynamic path length, $d(t)$, of the signals. The overall change of the signal can be described by $H(f,t)$, as shown in Fig.1, and can be expressed as:

$$\begin{aligned} H(f,t) &= H_s(f,t) + H_d(f,t) \\ &= H_s(f,t) + A(f,t)\, e^{-j2\pi \frac{d(t)}{\lambda}}, \end{aligned} \tag{2}$$

where $H_s(f,t)$ is the sum of static paths component and $H_d(f,t)$ represents the dynamic path component that varies with the dynamic path length $d(t)$; $A(f,t)\, e^{-j2\pi \frac{d(t)}{\lambda}}$ is the complex-valued representation of the attenuation and initial phase offset of $H_d(f,t)$ and λ is the wavelength of the carrier.

When the target moves a short distance, $A(f,t)$ can be considered as a constant [15], meaning that $|H_d(f,t)|$ can also be considered as a constant.

Fig. 2 shows the phase derivation of $H(f,t)$. The parameter p is the distance from the edge of $H_d(f,t)$ to $H_s(f,t)$; θ is the phase difference between $H_d(f,t)$ and $H_s(f,t)$; α is the phase of the $H(f,t)$; β is the phase of the $H_s(f,t)$; φ is the phase difference between $H_s(f,t)$ and

$$\begin{aligned} &H(f,t); \mu \\ &= \sqrt{|H_s(f,t)|^2 + |H_d(f,t)|^2 + 2|H_s(f,t)||H_d(f,t)|\cos\theta}. \end{aligned}$$

The phase of $H(f,t)\,\alpha$ can be described as:

$$\alpha = \beta - \varphi = \beta - acrsin\left(\frac{p}{|H(f,t)|}\right)$$

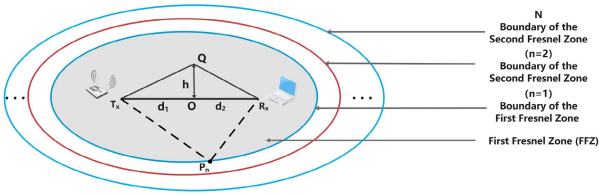


FIGURE 3. The model of Fresnel Zone. Q is the position of the human body and O denotes the vertical point from Q to T_xR_x.

$$= \angle H_s(f,t) - acrsin\left(\frac{|H_d(f,t)|sin\theta}{\mu}\right). \tag{3}$$

Since the $|H_d(f,t)|sin\theta$ is a small value compared to $H_s(f,t)$, the phase of $H(f,t)\,\alpha$ can be estimated to be as:

$$\alpha \approx \angle H_s(f,t) - \frac{|H_d(f,t)|}{\mu} sin\theta, \tag{4}$$

where the waveform of the CSI phase is a sinusoidal-like waveform. Furthermore, for the complementarity between phase and amplitude [12], the waveform of CSI amplitude is a cosinusoidal-like waveform.

B. FRESNEL ZONE MODEL

The Fresnel Zone (FZ) model was introduced in WiFi-based wireless sensing in [20], [21] and revealed the relationship between human activity and the received signals.

In the free space scenario, the transmitter Tx uses radio frequency (RF) signals with wavelength λ as the transmission medium to send information to the receiver Rx. The Fresnel Zones are concentric ellipses with foci of transmitter and receiver. The boundary of the *nth* Fresnel Zone is defined as:

$$|T_xP_n| + |R_xP_n| - |T_xR_x| = n\frac{\lambda}{2}. \tag{5}$$

The P_n is a point on the nth ellipse and the innermost ellipse ($n = 1$) is defined as the First Fresnel Zone (FFZ) as shown in Fig. 3.

[22] showed that between a pair of WiFi transceivers, more than 70% of the RF signal energy is transferred via the FFZ. Human activity will result in a path difference Δd of signal transmission.

The path difference Δd between $|T_xQR_x|$ and $|T_xOR_x|$ is defined as:

$$\begin{aligned} \Delta d &= |T_xQR_x| - |T_xOR_x| \\ &= \sqrt{d_1^2 + h^2} + \sqrt{d_2^2 + h^2} - (d_1 + d_2) \\ &= d_1\sqrt{1 + \left(\frac{h}{d_1}\right)^2} + d_2\sqrt{1 + \left(\frac{h}{d_2}\right)^2} - (d_1 + d_2), \end{aligned} \tag{6}$$

In FFZ, $\left(\frac{h}{d_1}\right)^2 \ll 1$ and $\left(\frac{h}{d_2}\right)^2 \ll 1$. Therefore, the path difference Δd can be approximated as:

$$\Delta d \approx \frac{h^2(d_1+d_2)}{2d_1d_2}. \tag{7}$$

The phase difference $\Delta\varphi$ caused by the path difference Δd can be expressed as:

$$\Delta\varphi = \frac{2\pi\Delta d}{\lambda} = \pi h^2 \frac{(d_1+d_2)}{\lambda d_1 d_2}. \quad (8)$$

When at the boundary of the FFZ, the phase difference $\Delta\varphi$ is π, but because of reflection, there is an offset of π in phase. Eventually, the two signals were out of phase by 2π, and the receiver received an enhanced signal. On the contrary, when at the boundary of the Second Fresnel Zone, the receiver received a weakened signal. Hence, when people do the same activity at different positions, the amplitude of the CSI signals may vary significantly [20], [21], [22], [23], [24]. In other words, activity recognition cannot be judged only by amplitude. The complementarity of amplitude and phase was used to solve the ''blind spot'' in recognition of amplitude or phase [12].

However, inherent to the hardware issues on the transmitter and receiver, there will be a random phase offset $e^{-j\cdot\theta_{offset}}$ during the CSI sampling process. The $e^{-j\cdot\theta_{offset}}$ represents the phase error introduced by central frequency offset (CFO), phase locked loop (PLL), sample frequency offset (SFO), and packet boundary detection (PBD). Several solutions have been proposed to recover the phase of CSI from commodity WiFi [24], [25], [26], [27], [28]. The above phase recovery methods require the network interface controller (NIC) to have at least two antennas and apply conjugate multiplication of CSI between two antennas to remove the time-varying random phase offset. However, Raspberry PI has only one antenna and cannot eliminate the time-varying phase offset by conjugation, so this method is not applicable for Raspberry PI.

The phase received by the receiver is described as:

$$\hat{\varphi}_i = \varphi_i - 2\pi\frac{k_i}{N}\delta + \beta + Z, \quad (9)$$

where φ_i is the actual phase; $\hat{\varphi}_i$ is the measurement phase; k_i is the index of the current subcarrier; N is the number of points in the Fast Fourier transform; β is a random phase offset caused by equipment; Z is the random noise in the measurement. [29] adopted a linear approach to eliminate the errors $2\pi k_i\delta/N$ and β.

$$a = \frac{\hat{\varphi}_n - \hat{\varphi}_i}{k_n - k_1} = \frac{\varphi_n - \varphi_i}{k_n - k_1} - \frac{2\pi}{N}\delta,$$
$$b = \frac{1}{n}\sum_{j=1}^{n}\hat{\varphi}_j = \frac{1}{n}\sum_{j=1}^{n}\varphi_j - \frac{2\pi\delta}{nN}\sum_{j=1}^{n}k_j + \beta,$$
$$\tilde{\varphi}_i = \hat{\varphi}_i - ak_i - b = \varphi_i - \frac{\varphi_n - \varphi_i}{k_n - k_1}k_i - \frac{1}{n}\sum_{j=1}^{n}\varphi_j, \quad (10)$$

where $\tilde{\varphi}_i$ is the value after linear transformation. Nevertheless, this method ignores the effects of random noise, where a filter is required to filter it out.

C. THE EXTRACTION OF OPTIMAL PHASE MATCHING STRATEGY

The existence of the pilot carrier, empty carrier, and guard carrier make not all phases correlated with activities [11], and the phase of some carriers may be seriously disturbed. All these make it challenging to extract the correct phase value. Therefore, it is critical to extract the optimal phase. The amplitude and phase of the WiFi signal are orthogonal, and the amplitude and phase received at the receiver are also nearly orthogonal. According to this characteristic, this paper used the amplitude value to derive and extract the optimal phase. However, amplitude sequence and phase sequence of CSI are not synchronized. The (DTW) algorithm can solve the problem and test the similarity of the two time series. The smaller the value is, the more similar the two time series are.

The amplitude sequence as a reference template is described as:

$$A(M) = \{a_1, a_2, \cdots, a_M\}, \quad (11)$$

where $a_1, a_2, \cdots, a_M$ are the corresponding value of each sampling point, and M is the total number of the points sampled.

The *nth* phase sequence as a test template is described as:

$$P_n(M) = \{P_{n,1}, P_{n,2}, \cdots, P_{n,m}\}. \quad (12)$$

From all phase carriers $P(N)$ to find and the most similar phase $P_n(M)$, where $P(N) = \{P_1(M), P_2(M), \cdots, P_N(M)\}$, N is the number of carriers collected.

The Z-score was adopted for both $A(M)$ and $P_n(M)$ to eliminate the phase differences between carriers and reduce the complexity of the calculation. Euclidean distance was used to calculate the distance between two sequence data points. The Euclidean distance between the *ith* data a_i of the amplitude sequence and the *j*th data p_j of the phase sequence $P_n(M)$ is calculated as shown:

$$d(a_i, p_j) = \sqrt{(a_i - p_j)^2}, \quad (13)$$

which produce a $M \times M$ distance matrix D. A path from $D_{1,1}$ to $D_{M,M}$ was determined where the sum of the elements on the path is the smallest, and it should obey the following restrictions: firstly, the starting point and the ending point of two sequences are consistent. secondly, the path must satisfy the chronological sequence of the data. The shortest path length is defined as $L(i, j)$.

$$L(i,j) = min\begin{cases} L(i-1,j) + D(i,j), \\ L(i-1,j-1) + 2D(i,j) \\ L(i,j-1) + D(i,j). \end{cases} \quad (14)$$

The shortest path of the current element must be the length of the shortest from the previous element plus the value of the current element, and the previous element must be selected from the adjacent to left as shown in (14). Traverse each in carrier phase $P(N)$ and find the shortest path $L_{min}(M, M)$.

The optimal phase matching strategy is summarized as follows:

Algorithm 1 Optimal Phase Matching Strategy

1) Initialization
- Input $A(M)$ and $P(N)$. The $P_n(M) \in P(N)$;$P_j \in P_n(M)$; $a_i \in A(M)$;
- Create an array O to hold each $P_n(M) \in P(N)$ distance from $A(M)$;

2) For all $P_n(M) \in P(N)$

Calculate the distance between all $a_i \in A(M)$ and $P_j \in P_n(M)$ and store the values in an $M \times M$ matrix D;

Create a new $M \times M$ matrix L and generate each element in L by using the strategy of (14). Store the value of $L(M, M)$ in O;

End for

3) Find the smallest value in O.

The carrier phase corresponding to this value is the most similar phase $P_s(M)$.

Output: $P_s(M)$

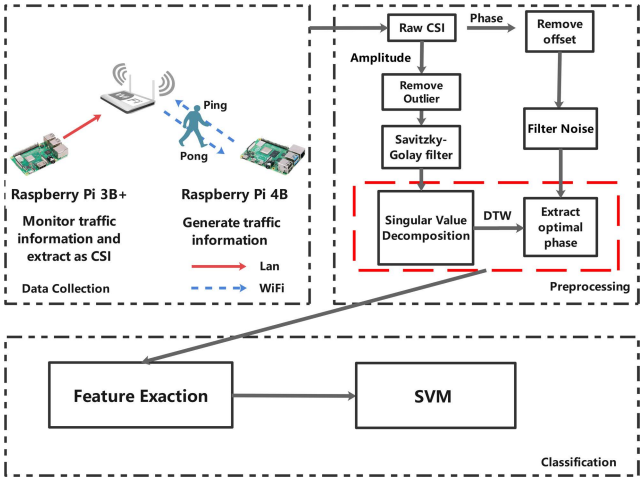


FIGURE 4. The system structure of wireless sensing including data collection and preprocessing and activity recognition.

IV. SYSTEM DESIGN

A wireless sensing system for human activity was designed in this paper, which the optimal phase was extracted for recognition based on the property of orthogonality between amplitude and phase. Fig. 4 shows the architecture of the wireless sensing system. The whole system includes three parts: data collection, preprocessing, and classification.

A. DATA COLLECTION

There were two Raspberry PI devices in the system, but they have their own roles. The Raspberry 3B+ with Nexmon patched was used to monitor and extract the CSI from WiFi signal. And that the Raspberry 4B was used to generated the traffic over so that the activity can be sampled at the same frequency. The TP-link AX3000 as the access point (AP) was set to IEEE 802.11 ac mode, bandwidth was 80MHz, the frequency was 5.21GHz, and the channel number was 157. Table. 1 shows the device configuration in the network.

TABLE 1. Device configuration in the network.

Name	Connection Method	Description
Router: TP-link AX3000	WLAN	Build the network and connect devices to the network
Raspberry PI 4B	WLAN	Create Traffic over WiFi
Raspberry PI 3B+	LAN	Monitor the router and extract as CSI
CSI		Sent from 10.10.10.10 to 255.255.255.255

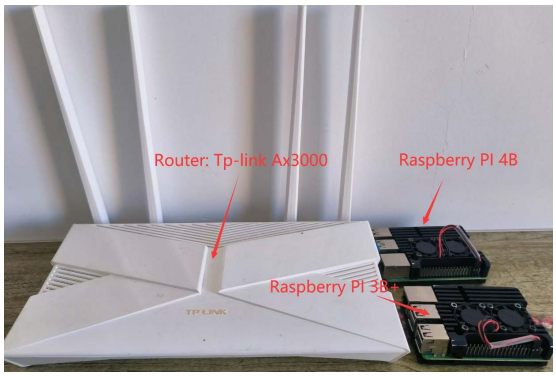


FIGURE 5. The devices of the wireless sensing system: Raspberry PI 3B+; Raspberry PI 4B; router.

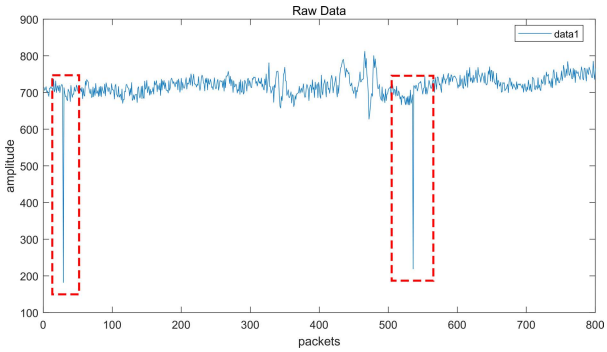


FIGURE 6. The raw data which contains plenty of noise and outliers.

B. PREPROCESSING

The noise in the wireless channel will interfere with the collected data. Therefore, it is necessary to preprocess the data to ensure accuracy. Fig. 6 shows the raw data where it be observed that the marked boxes were not caused by human activity but simply by sudden burst of noise.

The Savitzky-Golay [33] filter was used to denoise the data. The most important feature of this filter is to eliminate noise while ensuring the shape of the signal. Fig. 7 shows the filtered amplitude value.

In this system, the sampling frequency of CSI was 100 packets per second and the time of human activity was about 0.5 seconds to 2 seconds. Therefore, a sliding window

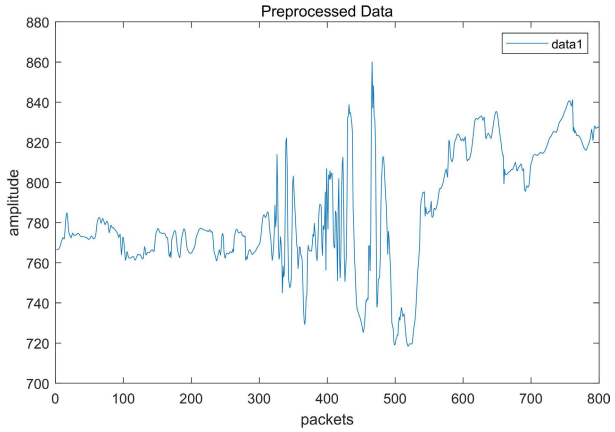


FIGURE 7. The filtered data of amplitude.

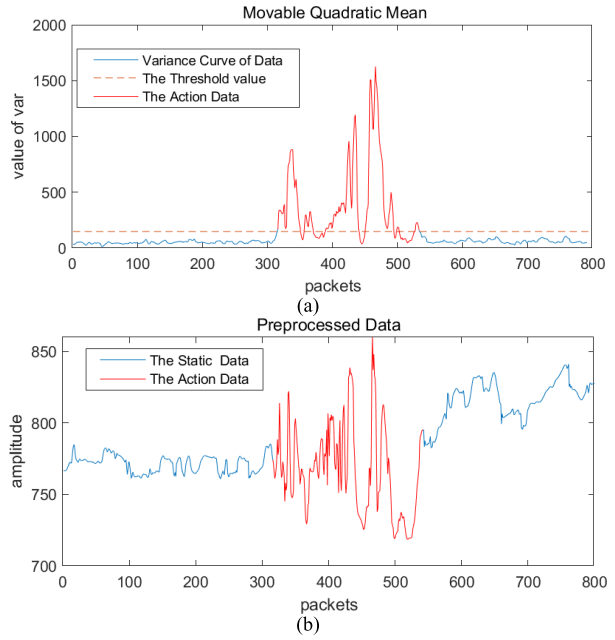


FIGURE 8. The segmentation of dynamic and static parts. (a) shows the use of variance to divide the data into dynamic and static parts. (b) shows the dynamic part of activity sampling.

with a length of 50 was designed to calculate the variance of CSI amplitude. In the whole activity, the dynamic part only took up a small part of it. Hence the average variance value was used as the threshold to distinguish the active part from the static part as shown in Fig. 8(a). The advantage of this was to extract the activity data, which greatly reduces the amount of data that needs to be processed. The activity curve was divided into static parts (blue) and dynamic parts (red) as shown in Fig.8(b).

In terms of the phase of CSI, there was a phase offset due to the time synchronization problem between the transmitter and receiver. A linear relationship was applied to descramble the phase [29]. Fig.9 shows the principal component variance of CSI. The high proportion of the amplitude's first principal component indicates that each carrier's amplitude is similar. In contrast, the proportion of the first principal component of the phase is not as high, which means that the phase on

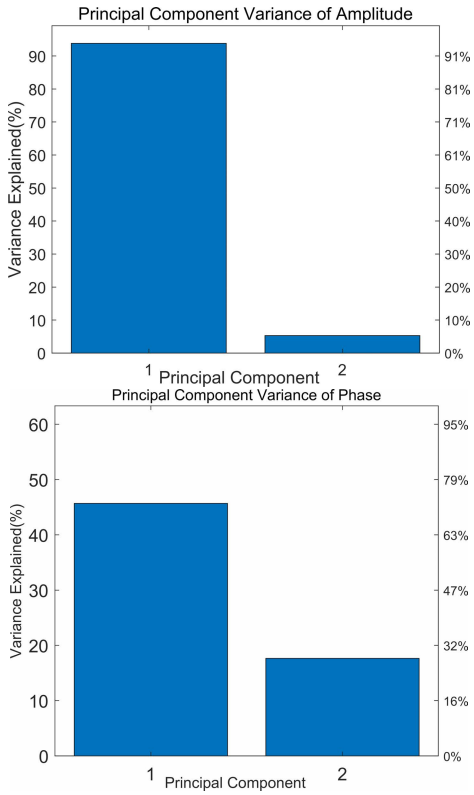


FIGURE 9. The principal component variance of CSI.

some carriers was seriously disturbed so that the phase on each carrier cannot be similar. Because of this, the phase after singular value decomposition cannot be directly used for activity recognition.

The DTW algorithm, described in section III C, was used to find the optimal phase value for activity recognition. The Fig. 10 (a) shows the original phase value, which includes much noise. Fig. 10 (b) shows an activity's amplitude and optimal phase values.

C. CLASSIFICATION

Different activities will have different impacts on in CSI. Therefore, in this paper, the characteristics of the activity curves were extracted for recognition, as shown in Table. 2.

These characteristics of activities and their corresponding name labels were taken as the input for the classification. The Support Vector Machine (SVM) can construct a multiclass classifier to divide samples of a specific category into one category and the rest of the samples into another category. The unknown samples were classified into the maximum classification function value category. The accuracy of the classification model was finally obtained by comparing it with the actual values.

V. EVALUATION

In this system, the following four types of activity were evaluated, as shown in Fig. 11. Three people performed each

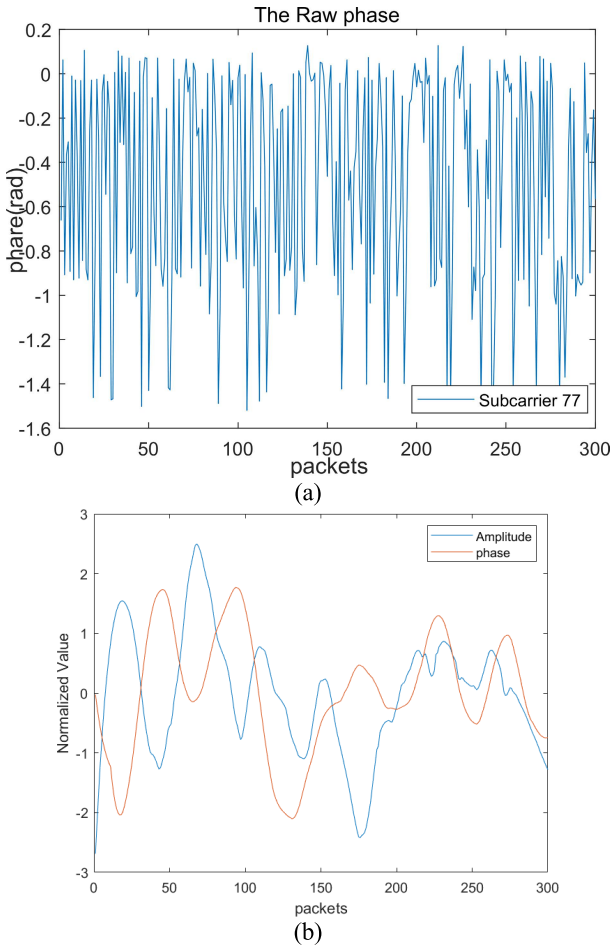


FIGURE 10. The processing of the phase. (a) is shown unscrambled phase value, it contains too much noise; (b) is shown the phase that best matches the amplitude, it filters out most of the noise and is orthogonal to the amplitude.

FIGURE 11. (a) stand (without any activity); (b) bend over; (c) squat; (d) hold up hands.

activity about 120 times to complete the activity data collection in 2 scenarios.

A. EXPERIMENT SETUP AND CONFIGURATION

Scenario 1 was in the 4 $m \times 6$ m conference room, where the router and the Raspberry PI 3B+ were 2.5 m apart. The layout of the environment is shown in Fig. 12.

Scenario 2 was in the 4 $m \times 10$ m office, where the router and the Raspberry PI 3B+ were 2.5 m apart. The layout of the environment is shown in Fig. 13.

TABLE 2. The Definition and formula of characteristics.

Characteristic	Definition	Formula
Arithmetic Mean	A measure of the amount of variation or dispersion of a set of values.	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
Max	A maximum in the data	$\{x_{max} \epsilon x_i \mid x_i \leq x_{max}\}$
Standard Deviation	A measure of the amount of variation or dispersion of a set of values.	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$
Median Absolute Deviation	The average of the absolute deviations from a central point	$\frac{1}{N}\sum_{i=1}^{N}\lvert x_i - m(X)\rvert$
Interval Range	A set of real numbers that contains all real numbers lying between any two numbers of the set.	$\{x \epsilon x_i \mid a \leq x \leq B\}$
Interquartile Range	A measure of statistical dispersion, which is the spread of the data.	$IRQ = Q_3 - Q_1$

In this paper, the scenario 2 was defined as a non-ideal case where plenty of wireless devices and obstacles were considered. The Wireless devices will affect wireless transmission channels and signals. Plenty of obstacles will affect the transmission path of the signal.

B. EXPERIMENTAL RESULTS

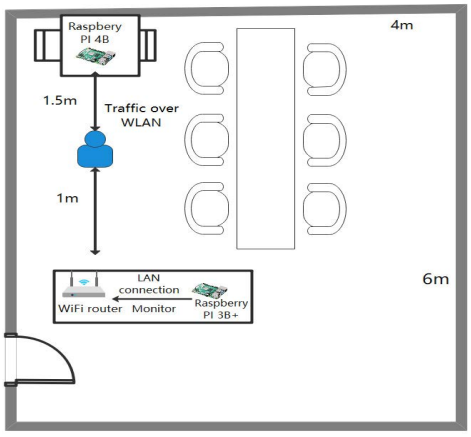
In this paper, 80% of the collected data was used as the training set to train the classification model, and the remaining data was used as the test set to measure the performance of the classification model.

The Fig. 14 shows the performance of the classification model in the conference room environment. The method proposed in this paper can improve the accuracy of recognition. In terms of holding up hands, the classification model that only used phase and the model only used amplitude both performed poorly. On the contrary, the model using the amplitude and optimal phase had excellent classification performance. In other actions, the accuracy of the proposed method was slightly higher than other methods.

In addition, this paper compares the accuracy of extracting all phase values and extracting the best value. As shown in the yellow and blue bars, the performance of extracting all phase values was very poor, because all phase values contained much noise, resulting in a deviation from the real value.

In scenario 2, there were many network devices that interfere with CSI collection and transmission. However, the accuracy of using the optimal phase proposed in this paper was

(a)



(b)

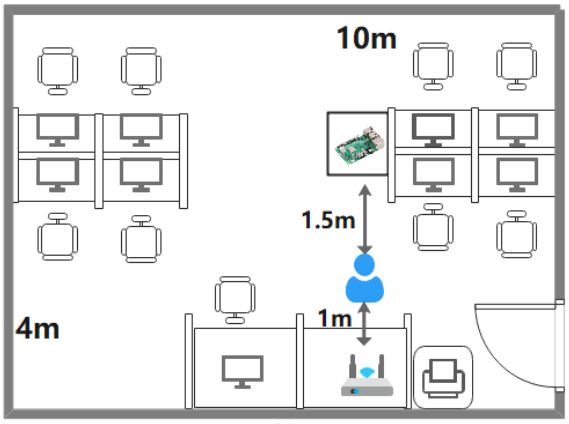
FIGURE 12. The layout of the scenario 1: Conference Room. (a) The test scenario diagram in real life. (b) The overall layout of the test environment.

still much higher than that of using all phases. Fig. 15 shows the performance of the classification model in the office environment.

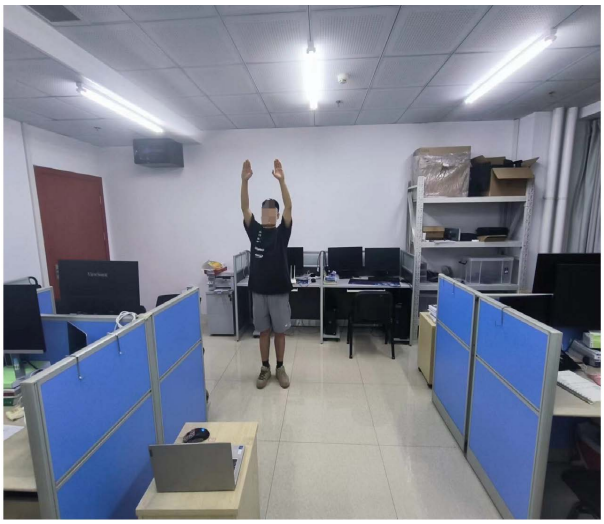
Fig. 16 shows the true positive for the two scenarios. The accuracy of all actions decreased compared to scenario 1. This was because the received signal in scenario 2 was more diffracted and its dynamic component was more complex. Moreover, the wireless device will also affect the CSI. It can be seen that the method of extracting the optimal phase proposed in this paper is superior to other methods, whether in a wireless environment with fewer devices or a relatively more complex environment.

Therefore, it can be concluded that it is crucial to use the method in this paper to find the optimal phase, which can help us improve the accuracy of the system. Furthermore, no matter what the scenario, the static state had a high accuracy, which makes it easy to distinguish the dynamic state from the static state.

This paper compared the recognition methods of other articles. In [10], a WiFi-based system was also used for human activity recognition. Six specific actions were identified by collecting the amplitude information of CSI. In [30], the deep neural network was used to recognize 6 types of actions through plenty of experiments. However, only the amplitude information of CSI was used. In [13], the Intel 5300 was used to collect CSI, and the local outlier factor of amplitude was



(a)



(b)

FIGURE 13. The layout of the scenario 2: Office. (a) The test scenario diagram in real life. (b) The overall layout of the test environment.

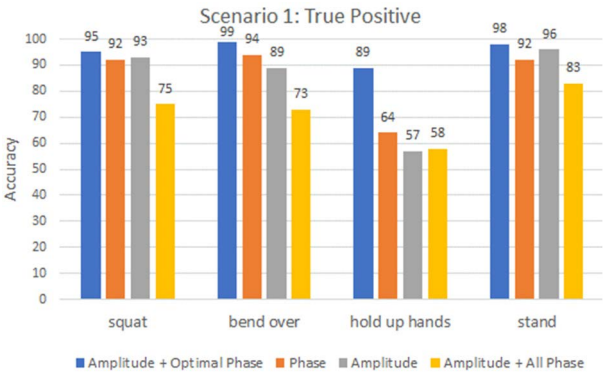


FIGURE 14. Classification Model Performance for scenario 1: Conference Room.

used to judge. The above three research works are only using the amplitude to recognize, which wastes the extracted phase information. The optimal phase extraction method proposed in this paper can help increase the system gain and improve the performance of system recognition Table. 3 shows the recognition accuracy of different research works.

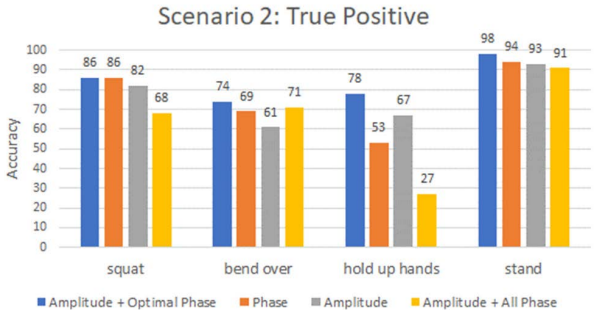


FIGURE 15. Classification Model Performance for scenario 2: Office.

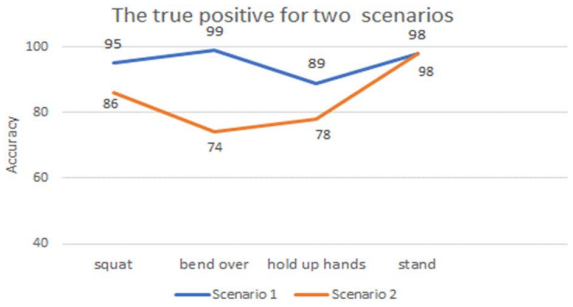


FIGURE 16. The true positive for the two scenarios.

TABLE 3. The recognition accuracy of different research works.

Methods	Overall Accuracy
This paper proposed	95%
Study [10]	90%
Study [30]	97%
Study [13]	90%

VI. CONCLUSION

In the wireless recognition system, the phase was sensitive to noise, resulting in a deviation in the acquired phase value and a decrease in accuracy. In this paper, a method to find and extract the optimal phase was proposed. This method found the value most similar to the amplitude from all the phase values according to the orthogonality of the amplitude and phase. Furthermore, this paper also compared the performance of using this method and not using this method in different scenarios. The estimation results showed that the proposed method could improve the accuracy of the recognition strategy with commercial WiFi signals.

In future work, we will focus on extracting more information from the phases, such as the speed and direction of the activities. This can better reshape the motion to achieve the goal of integrated sensing and communication.

ACKNOWLEDGMENT

The authors would like to thank for the contribution of Xiamen Sage Tech. Ltd. in the on-going development of the Metaverse Engine Design.

REFERENCES

[1] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf, "Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features," in *Proc. Int. Conf. Adv. Mobile Comput. Multimedia (MoMM)*, Vienna, Austria, 2013, pp. 43–52.
[2] H. Abdelnasser, M. Youssef, and D. K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr./May 2015, pp. 1472–1480.
[3] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–32, 2013.
[4] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Chennai India, Jul. 2017, pp. 1–10.
[5] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using WiFi channel state information," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 98–104, Oct. 2017.
[6] F. Zhang, F. Zhang, Z. Chang, and D. Zhang, "From Fresnel diffraction model to fine-grained human respiration sensing with commodity WiFi devices," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–23, Mar. 2018.
[7] D. Zhang, D. Wu, K. Niu, X. Wang, F. Zhang, J. Yao, D. Jiang, and F. Qin, "Practical issues and challenges in CSI-based integrated sensing and communication," 2022, *arXiv:2204.03535*.
[8] M. Schulz et al. "Demo: Nexmon in action: Advanced applications powered by the nexmon firmware patching framework," in *Proc. 11th ACM Int. Workshop Wireless Netw. Testbeds, Exp. Eval. Characterization*, Snowbird, UT, USA, 2017.
[9] W. Zhang, Z. Wang, and X. Wu, "WiFi signal-based gesture recognition using federated parameter-matched aggregation," *Sensors*, vol. 22, no. 6, p. 2349, Mar. 2022.
[10] L. Tao, C. Shi, P. Li, and P. Chen, "A novel gesture recognition system based on CSI extracted from a smartphone with nexmon firmware," *Sensors*, vol. 21, no. 1, p. 222, 2020.
[11] M. S. Gast, "The PHY," in *802.11ac: A Survival Guide*. Sebastopol, CA, USA: O'Reilly Media, 2013, pp. 11–36.
[12] 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 Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–19, Sep. 2018.
[13] 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.
[14] S. Tan and J. Yang, "WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Paderborn, Germany, Jul. 2016, pp. 201–210.
[15] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "Farsense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas," in *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–26, 2019.
[16] 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.*, Denver, CO, USA, May 2017, pp. 1961–1972.
[17] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single Wi-Fi link," in *Proc. 16th Annu. Int. Conf. Mobile Syst., Appl., Services*, Munich, Germany, Jun. 2018, pp. 350–361.
[18] R. Gao, M. Zhang, J. Zhang, Y. Li, E. Yi, D. Wu, L. Wang, and D. Zhang, "Towards position-independent sensing for gesture recognition with Wi-Fi," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 2, pp. 1–28, Jun. 2021.
[19] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, and Y. Liu, "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proc. 17th Annu. Int. Conf. Mobile Syst., Appl., Services*, Seoul, South Korea, Jun. 2019, pp. 313–325.
[20] H. Wang, D. Zhang, J. Ma, Y. Wang, and Y. Wang, "Human respiration detection with commodity WiFi devices: Do user location and body orientation matter?" in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Berlin, Germany, Sep. 2016, pp. 25–36.
[21] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "WiDir: Walking direction estimation using wireless signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Berlin, Germany, Sep. 2016, pp. 351–362.
[22] D. H. Hristo, *Fresnel Zones in Wireless Links, Zone Plate Lenses and Antennas*. Boston, MA, USA: Artech House, 2000.
[23] F. Zhang, K. Niu, J. Xiong, B. Jin, T. Gu, Y. Jiang, and D. Zhang, "Towards a diffraction-based sensing approach on human activity recognition," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 1, pp. 1–25, Mar. 2019.
[24] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with Wi-Fi signals," *Computer*, vol. 50, no. 1, pp. 48–57, Jan. 2017.

[25] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter level localization using WiFi," in *Proc. ACM Conf. Special Interest Group Data Commun.*, London, U.K., Aug. 2015, pp. 269–282.
[26] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-MUSIC: Accurate device-free indoor localization," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Berlin, Germany, Sep. 2016, pp. 196–207.
[27] N. Yu, W. Wang, A. X. Liu, and L. Kong, "QGesture: Quantifying gesture distance and direction with WiFi signals," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–23, Mar. 2018.
[28] J. Zhu, Y. Im, S. Mishra, and S. Ha, "Calibrating time-variant, device-specific phase noise for COTS WiFi devices," in *Proc. 15th ACM Conf. Embedded Netw. Sensor Syst.*, Delft, The Netherlands, Nov. 2017, pp. 1–12.
[29] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "PhaseU: Real-time LOS identification with WiFi," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Hong Kong, Apr. 2015, pp. 2038–2046.
[30] J. Schäfer, B. R. Barrsiwal, M. Kokhkharova, H. Adil, and J. Liebehenschel, "Human activity recognition using CSI information with nexmon," *Appl. Sci.*, vol. 11, no. 19, p. 8860, Sep. 2021.
[31] S. Kim, S. Park, B. Na, J. Kim, and S. Yoon, "Towards fast and accurate object detection in bio-inspired spiking neural networks through Bayesian optimization," *IEEE Access*, vol. 9, pp. 2633–2643, 2021.
[32] Y. Zhang, Y. Zheng, G. Zhang, K. Qian, C. Qian, and Z. Yang, "GaitSense: Towards ubiquitous gait-based human identification with Wi-Fi," *ACM Trans. Sensor Netw.*, vol. 18, no. 1, pp. 1–24, Feb. 2022.
[33] P. A. Gorry, "General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method," *Anal. Chem.*, vol. 62, no. 6, pp. 570–573, 1990.

CHEN TIAN (Student Member, IEEE) received the bachelor's degree from the Huaiyin Institute of Technology, in 2020. He is currently pursuing the master's degree in opto-electronic and communication engineering with the Xiamen University of Technology. His research interest includes 5G/6G communication system techniques, such as contactless wireless sensing.

YUE TIAN (Member, IEEE) received the bachelor's degree from the Beijing University of Posts and Telecommunications, in 2007, and the master's degree in wireless communication and signal processing and the Ph.D. degree in electrical engineering from the University of Bristol, in 2011 and 2017, respectively. In 2017, he joined the School of Opto-Electronic and Communication Engineering, Xiamen University of Technology, as an Associate Professor. His research interests include 5G/6G communication system techniques, such as massive MIMO, non-orthogonal multiple access, and UAV communication networks.

XIANLING WANG (Member, IEEE) received the B.S. degree in communication engineering and the Ph.D. degree in communication and information system from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China, in 2009 and 2014, respectively. He is currently an Assistant Professor with the Fujian Key Laboratory of Communication Network and Information Processing, Xiamen University of Technology, China. He has authored or coauthored several papers in international conferences and journals in his research areas. His research interests include stochastic geometry, game theory, 5G wireless networks, UAV communication networks, and nonorthogonal multiple access.

YAU HEE KHO (Senior Member, IEEE) was born in Kuching, Malaysia, in 1974. He received the B.Eng. (Hons.) and Ph.D. degrees from the University of Canterbury, Christchurch, New Zealand, in 1997 and 2008, respectively, and the Professional Certificate in learning and teaching (higher education) from the Swinburne University of Technology, Melbourne, Australia, in 2013. From 1998 to 2002, he was an Electrical Design Engineer in Singapore. He was a Senior Lecturer at the Swinburne University of Technology (Sarawak Campus), from 2009 to 2013, and was an Assistant Professor at Nazarbayev University, from 2013 to 2017. He is currently a Senior Lecturer and the Programme Director with the Victoria University of Wellington, New Zealand. His research interests include wireless communications, signal processing, electronics, and engineering education.

ZHENZHE ZHONG (Member, IEEE) received the undergraduate degree from BUPT, in 2013, the M.Sc. degree in wireless communication and signal processing from the University of Bristol, in 2014, and the doctor degree from Télécom Paris (a School of Institut Polytechnique de Paris), in 2020.

He is the Wireless Communications Expert (Research and Development Manager) with Xiamen Intretech Inc. and a Joint Programme Post-doctoral Research Fellow with Tsinghua University. He worked in the Toshiba Telecommunication Research Laboratory as a Research Engineer, from 2014 to 2016. He worked in Orange Laboratories, Lannion, France, as a Ph.D. Research Engineer Candidate, from 2016 to 2019. He has published over ten research papers and granted three patents. He has also published over 20 patents in his latest Research and Development job. His research interests include wireless communications, congestion control, mobile edge/fog computing, smart manufacturing, digital twins, and machine learning.

WENDA LI (Member, IEEE) received the M.Eng. and Ph.D. degrees in electrical and electronic engineering from the University of Bristol, in 2013 and 2017, respectively. He was at the University of Birmingham as a Research Fellow. In 2019, he joined University College London, where he is currently working as a Research Fellow with the Department of Security and Crime Science. His research in passive WiFi radar has led to a number of IEEE conference and journal publications. His research interests include signal processing for passive radar and high-speed digital system design for wireless sensing applications in healthcare, security, and positioning.

BAIYUN XIAO (Student Member, IEEE) received the bachelor's degree from the Hunan Institute of Science and Technology, Yueyang, Hunan, China, in 2020. She is currently pursuing the master's degree in opto-electronic and communication engineering with the Xiamen University of Technology. Her research interest includes S5G/6G cooperative communication techniques, such as reconfigurable intelligent surfaces.