

# TW-See: Human Activity Recognition Through the Wall With Commodity Wi-Fi Devices

Xuangou Wu , Zhaobin Chu, Panlong Yang, Chaocan Xiang , Xiao Zheng, and Wenchao Huang 

**Abstract**—Device-free passive human activity recognition plays an important role in many applications, such as smart homes, identification, health care, etc. However, existing human activity recognition systems either require a dedicated device or do not meet the scenarios of the signals through the wall. To address this challenge, we present TW-See, a device-free passive human activity recognition system with Wi-Fi signals, which does not require any dedicated device and meets the scenarios of the signals through the wall. TW-See mainly exploits two key techniques to recognize different human activities. First, we propose an opposite robust PCA (Or-PCA) approach to obtain the correlation between human activity and its resulting changes in channel state information values that can eliminate the influence of the background environment on correlation extraction. Second, we propose a normalized variance sliding windows algorithm to segment the human action time from the Or-PCA waveforms, which can distinguish the true or false of the start and end times of human actions. Furthermore, we implemented TW-See with commodity Wi-Fi devices and evaluated it in several different environments. Experimental results show that TW-See achieves an average accuracy of 94.46% when the signals pass through the concrete wall.

**Index Terms**—Human activity recognition, CSI, device-free, commodity Wi-Fi.

## I. INTRODUCTION

HUMAN activity recognition plays an important role in pervasive computing and human-computer interaction, and can support many emerging applications such as smart homes, augmented reality, identification, health care, etc. Many human activity recognition systems have been proposed with different techniques, such as wearable sensor based approaches

Manuscript received March 28, 2018; revised June 17, 2018 and September 27, 2018; accepted October 27, 2018. Date of publication October 30, 2018; date of current version January 15, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants 61672038, 61872447, 61772546, 61502520, and 61502010, in part by the Visiting Study Project for Outstanding Young Talents of Anhui under Project GXGWFX2018017, in part by the Major Technologies R&D Special Program of Anhui under Grant 16030901060, and in part by the National Science Foundation of Jiangsu For Distinguished Young Scientists under Grant BK20150030. The review of this paper was coordinated by Dr. S. Kim. (*Corresponding author: Xiao Zheng.*)

X. Wu, Z. Chu, and X. Zheng are with the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243002, China (e-mail: wxgou@mail.ustc.edu.cn; zhaobinchu.go@gmail.com; xzheng@seu.edu.cn).

P. Yang and W. Huang are with the School of Computer Science and Technology, University of Science and Technology of China, Hefei 230022, China (e-mail: plyang@ustc.edu.cn; huangwc@ustc.edu.cn).

C. Xiang is with the College of Computer Science, Chongqing University, Chongqing 400044, China, and also with the Army Logistics University of PLA, Chongqing 401311, China (e-mail: xiang.chaocan@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2018.2878754

(e.g., [1]–[3]), computer vision based approaches (e.g., [4]–[6]), ambient device based approaches (e.g., [7]–[9]) and so on. Wearable sensor based approaches are intended to recognize human activities by using sensors [1], RFID [2], smartphone [3], and so on. These systems use active detection techniques and require the device always on the body. Computer vision based approaches exploit the camera to capture image sequences, and recognize human activities using activity classification algorithms. However, the camera is affected by the illumination of lights (e.g., a dark room) and shelters (e.g., wall). Ambient device based approaches attempt to recognize human activities by using radar [7], infrared [8], audio [9] and so on. Although these techniques do not require wearing or carrying anything on the body, some dedicated devices should be deployed.

Recently, Wi-Fi signals based human activity recognition techniques have been widely studied, they exploit Channel State Information (CSI) and multi-antennas techniques to distinguish different activities, such as Wi-Vi [10], WiSee [11], E-eyes [12], CARM [13], WifiU [14], RT-Fall [15] and others. Wi-Vi [10] used Wi-Fi signals through the wall to detect objects and human actions, but it was implemented by USRP 210 instead of the commercial Wi-Fi devices. E-eyes [12] utilized the commercial Wi-Fi devices to monitor human activity with one transmitter and three receivers, which achieved in-place and walking activity recognition. CARM [13] exploited a CSI-speed model and a CSI-activity model to recognize different activities. WifiU [14] used CSI values to recognize humans by capturing the gait patterns of different humans. RT-Fall [15] exploited both the amplitude and phase of CSI measurements to detect falls. Although these systems have the advantages of passive detection and easy deployment, they did not consider human activity recognition under the scenarios where the Wi-Fi signals pass through the wall.

In fact, most of the indoor environments have multiple rooms that access one wireless signals Access Point (AP), especially in the home environments. In these scenarios, Wi-Fi signals must pass through the wall to be received. The wall may block all the direct and reflected propagation paths between the receiver and the transmitter. Experimental results show that Wi-Fi signals would be affected severely by the wall. For example, the Wi-Fi signals would be attenuated 18 dB when 2.4 GHz Wi-Fi signals pass through a concrete wall with 18 inches width [16]. Meanwhile, our experimental results show that the existing denoising techniques (e.g., low-pass filter, PCA) could result in arbitrarily bad activity recognition performance, when the distance between transmitter and receiver is over 3 m and all the

propagation paths between them are blocked by the concrete wall.

In this paper, we propose TW-See, a device-free passive human activity recognition system with Wi-Fi signals, which meets the scenarios of the Wi-Fi signals through the wall. TW-See consists of the most common homes/offices Wi-Fi devices (with one antenna) and any PC/laptop (with three antennas NIC). The homes/offices Wi-Fi devices act as the transmitter, and the PC/laptop act as the receiver. To realize TW-See system, we mainly face two technical challenges. The first technical challenge is to obtain useful human activity CSI correlation from the raw CSI measurements. It is a challenging problem to obtain the correlation because the Wi-Fi signals are affected severely by the wall and the indoor physical environment (e.g., reflection, diffraction, and scattering). These two aspects cause the received signal not only to become weak while mixing a large amount of complicated background environmental information. The second technical challenge is to segment activity from CSI waveforms. Existing work displays that the activity could be segmented easily. The reason is that the amplitude of received signal changes dramatically, when the signals between the transmitter and receiver have direct or reflected propagation paths. However, the changes of CSI waveform caused by some human activities are not noticeable, when the Wi-Fi signals pass through the wall.

To overcome the first challenge, we propose an Opposite robust PCA (Or-PCA) approach to obtain the correlation between human activity and its resulting changes in CSI values. Although many existing efforts obtain this correlation from the raw CSI measurements directly (e.g., [13], [15]), this correlation cannot reflect the real relationship between human activity and its resulting changes in CSI values. The reason is that this correlation includes the correlation of the indoor background environment and noise. If the correlation between human activity and its resulting changes in CSI values is weakly affected by the environment, it can be obtained directly from the raw CSI measurements. Otherwise, it cannot be obtained from the raw CSI measurements, especially in the scenarios of the signals through the wall. This is why existing work can extract this correlation directly on the raw CSI measurements, when the transmitter and receiver have the direct or reflected signal propagation paths. At the same time, they require selecting waveforms from the second and subsequent principal components to extract the correlation. Unlike existing work, we only exploit the changed CSI caused by the human activity to recognize human activities. It helps us to reduce the interference of the complicated background environment CSI and noise, and make the correlation concentrate on the first Or-PCA component. To overcome the second challenge, we propose a normalized variance sliding windows algorithm to segment human activities from the Or-PCA waveforms. Our algorithm can eliminate the influence of slight fluctuations of the Or-PCA waveforms on the human activity segmentation, and distinguish the true or false of the start and end time of human activity.

The contributions of our work are summarized as follows:

- We propose a device-free human activity recognition system with Wi-Fi signals, TW-See, which does not require

any dedicated device and meets the scenarios of the signals through the wall. To the best of our knowledge, this is the first work to leverage CSI for human activity recognition through the wall with commodity Wi-Fi devices.

- To realize TW-See, we propose an Or-PCA approach to obtain the correlation between human activity and its resulting changes in CSI values. Unlike existing approaches, Or-PCA not only can effectively extract the correlation but also can make the correlation concentrate on the first principal component. According to the obtained correlation, we extract the activity features and use a BP neural network to recognize different human activities.
- We implemented TW-See system using the commercial Wi-Fi devices, with only one antenna at the transmitter, and three antennas at the receiver. Experiment results show that TW-See achieves an average accuracy of 94.46% when the signals pass through the concrete wall.

The rest of this paper is organized as follows. We present a literature review of existing work in Section II. Section III presents the preliminaries. Empirical analysis and system overview are given in Section IV and V, respectively. Wi-Fi signals processing is presented in Section VI. Or-PCA based CSI correlation extraction is proposed in Section VII. In Section VIII, we give human activity detection in detail. Section IX reports our experiment results, and make a conclusion in Section X.

## II. RELATED WORK

In this section, we summarize the related work on human activity recognition and CSI-based activity recognition with Wi-Fi signals.

**Human activity recognition:** Existing efforts on human activity recognition mainly includes wearable sensor based systems (e.g., [1]–[3]), computer vision based systems (e.g., [4]–[6]), ambient device based systems (e.g., [7]–[9]).

Wearable sensor based approaches intend to recognize human activities by using sensors [1], RFID [2], smartphone [3] and so on. Yatani K *et al.* [1] proposed BodyScope, which exploits a wearable acoustic sensor to record the sounds and classify them into different activities (e.g., eating, drinking, speaking, laughing, and coughing). Fortin-Simard D *et al.* [2] proposed a new affordable activity-recognition system based on passive RFID technology, which can detect errors related to cognitive impairment. Khan A M *et al.* [17] used a smartphone with a built-in triaxial accelerator for human activity recognition. These approaches are active detection techniques, which require the device always on the body.

Computer vision based approaches exploit the camera to capture image sequences, and recognize human activities using activity classification algorithms. G. Debard *et al.* [4] proposed a Camera-based fall detection system, which is based on background subtraction and measurements extracted from the dominant foreground object. H Foroughi *et al.* [5] proposed a novel method to detect various posture-based events, which exploits video surveillance to monitor old people in the home environment. However, the camera is affected by the illumination of

lights (e.g., a dark room) and shelters (e.g., wall). Besides, it also causes security and privacy issues.

Ambient device based approaches attempt to detect falls by exploiting the ambient information including infrared [8], radar [7], audio noise [9] and so on. S Tao *et al.* [8] exploits an infrared ceiling sensor network to realize behavior analysis and fall detection of a single person in the home environment. Q. Wu *et al.* [7] developed an effective human activity recognition scheme for the application in continuous-wave radar systems, which exploits time-frequency characteristics of the radar Doppler signatures, and the motion events are classified using the joint statistics of different features. Although these techniques do not require wearing or carrying anything on the body, some dedicated devices should be deployed in home environments.

**CSI-based activity recognition:** In recent years, many CSI-based systems were proposed to recognize human activities (e.g., Wi-Vi [10], WiSee [11], E-eyes [12], CARM [13], WifiU [14], RT-Fall [15], etc). These existing human activity recognition systems either require the dedicated device or do not meet the scenarios of the signals through the wall. For example, Wi-Vi [10] used Wi-Fi signals through the wall to detect objects and human actions, but it was implemented by USRP 210 instead of the commercial WiFi devices. E-eyes [12] utilized the commercial WiFi devices to monitor human activity with one transmitter and three receivers, which achieved in-place and walking activity recognition. CARM [13] exploited a CSI-speed model and a CSI-activity model to recognize different activities. WifiU [14] used CSI values to recognize humans by capturing the gait patterns of different humans. RT-Fall [15] exploited both the amplitude and phase of CSI measurements to detect falls.

Although these systems have the advantages of passive detection and easy deployment, they did not take into account human activity recognition under the scenarios where the signals pass through the wall. Indeed, most of the indoor environments have multiple rooms that access one wireless signals AP, especially in the home environment. In these scenarios, the wall may block all the direct and reflected propagation paths between the receiver and the transmitter.

### III. PRELIMINARIES

#### A. CSI and Propagation Model

In wireless communications, CSI describes how a signal propagates from the transmitter to the receiver. Meanwhile, CSI would be affected by the physical environment (e.g., reflection, diffraction and scattering). In a narrow band flat-fading channel with multiple transmission and receiving antennas (MIMO), the channel can be modeled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where  $\mathbf{x}$  and  $\mathbf{y}$  represent the transmitted and received vector, respectively.  $\mathbf{H}$  is the channel matrix and  $\mathbf{n}$  is the noise vector.

Due to the fact that the commercial Wi-Fi device uses the Orthogonal Frequency Division Multiplexing (OFDM) system,

CSI of the  $k$ -th subcarrier can be denoted as

$$H_{\text{eff}}(k) = |H_{\text{eff}}|e^{j \sin \theta} \quad (2)$$

where  $|H_{\text{eff}}|$  and  $\theta$  are the amplitude and phase of the subcarrier, respectively.

According to the free space propagation model [18], the Wi-Fi signals propagation model in the indoor environment can be denoted as

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \quad (3)$$

where  $P_r(d)$  is the received power,  $P_t$  is the transmitted power,  $G_r$  is the receiver antenna gain,  $G_t$  is the transmitter antenna gain,  $\lambda$  is the wavelength,  $d$  is the distance from transmitter to receiver, and  $L$  is system loss factor. In a typical indoor environment, considering the wall and human movements, the model can be represented as [19]:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 (d^2 + 4d_r^2 + \eta^2)L} - Lw \quad (4)$$

where  $d_r$  is the distance from reflection point to direct path,  $\eta$  is the change of path length caused by a moving human, and  $Lw$  is related to the material of the wall. For example,  $Lw$  equals to 18 dB if the wall is concrete with 18 inches width [16].

#### B. Fundamentals of Low-Rank Matrix Decomposition

Suppose that we have a matrix  $D = X + E$  ( $D \in \mathbb{R}^{n_1 \times n_2}$ ), where  $X$  is low-rank matrix and  $E$  is sparse errors. If we only know  $D$  or a part of  $D$ , how to recover  $X$  and  $E$ ? This problem is also called low-rank matrix completion problem. Recently, low-rank matrix recovery and completion via convex optimization has been widely studied (e.g., [20], [21]). Since the rank of  $X$  is unknown, the problem is to find a low-rank matrix that could have generated  $D$  with an unknown sparse matrix  $E$ . Therefore, it can be formulated as the following combinatorial optimization problem [20]

$$\begin{aligned} & \min \text{rank}(X) + \|E\|_0 \\ & \text{s.t. } D = X + E \end{aligned} \quad (5)$$

where  $\|\cdot\|_0$  is the counting norm (i.e., the number of non-zero entries in the matrix). Because  $\|\cdot\|_0$  is non-convex, the problem (5) is NP-hard problem. Then, the problem (5) could be relaxed as [21]

$$\begin{aligned} & \min \|X\|_* + \lambda \|E\|_1 \\ & \text{s.t. } D = X + E \end{aligned} \quad (6)$$

where  $\|X\|_*$  is the nuclear norm of matrix  $X$ , which is the sum of its singular values. That is,  $\|X\|_* = \sum_{i=1}^{\min\{n_1, n_2\}} \sigma_i$  and  $\sigma_i$  is the  $i$ -th singular value of  $X$ , where  $n_1$  and  $n_2$  are the length of the rows and columns of  $D$ , respectively.  $\lambda$  is a weighted factor. In fact, the problem (6) is also called as robust PCA, which is used to eliminate the impact of noise  $E$  on data  $D$ .

Many researches have been proposed to solve the problem (6), including iterative reweighted least squares algorithm [22], fixed point continuation with approximate SVD [23], the augmented

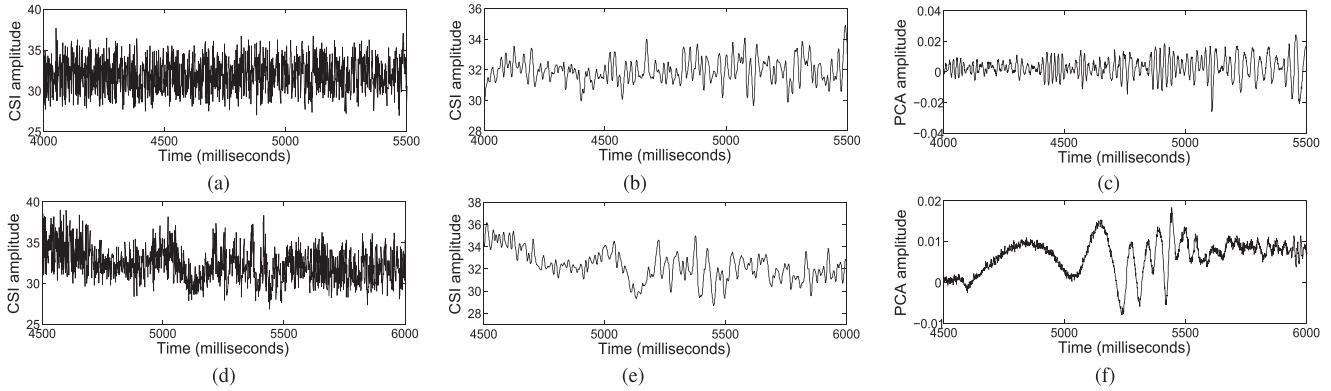


Fig. 1. Raw CSI and its denoising results when the transmitter and receiver are deployed in the same room. (a) Raw CSI of one subcarrier (walking). (b) Low-pass filter waveform (walking). (c) PCA waveform (walking). (d) Raw CSI of one subcarrier (falling). (e) Low-pass filter waveform (falling). (f) PCA waveform (falling).

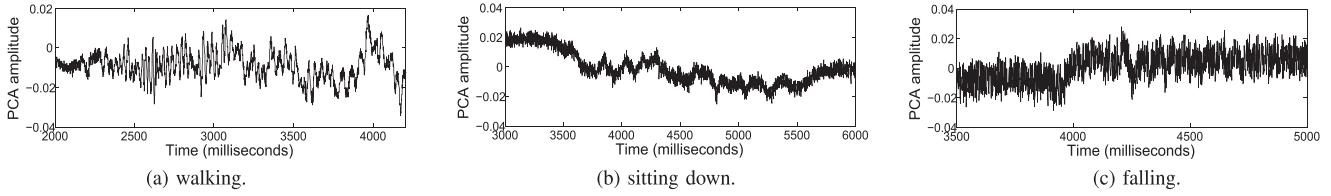


Fig. 2. PCA waveforms of different human activities when the Wi-Fi signals pass through the wall.

lagrange multiplier method [24], singular value thresholding [25] and so on.

#### IV. EMPIRICAL ANALYSIS

In this section, we give the empirical analysis of CSI based human activity recognition with different indoor environments. In our experiments, we use the Tenda W15E Router with one antenna as the transmitter, and use a Wi-Fi NIC (Intel 5300) with three antennas as the receiver. We collected the CSI measurements of 90 subcarriers because every antenna contains 30 subcarriers. We first give the existing denoising results of CSI amplitudes with different human actions. We then discuss why the CSI correlation cannot be extracted effectively using the existing denoising techniques when the Wi-Fi signals pass through the wall. Finally, we study the matrix characteristics of CSI measurements caused by the background environment and human activities.

##### A. CSI Measurement Denoising

First, we give the experimental results of existing denoising techniques (low-pass filter and PCA) when the transmitter and the receiver were deployed in the same room. In our experiment, the distance between the transmitter and the receiver was 3 m. Two human actions (walking and falling) occurred in the room, respectively. Since the transmitter and receiver were deployed in the same room, human action could interfere with the direct or reflected propagation paths of Wi-Fi signals. Fig. 1 shows the raw CSI of the two actions and the experimental results under different denoising techniques. Fig. 1(a) and (d) display the raw CSI amplitudes of one subcarrier.

Fig. 1(b) and (e) display low-pass filter waveforms of the raw CSI amplitudes, which show that CSI correlation cannot be extracted effectively. In our experiment, the band threshold of the low-pass filter is 100 Hz. Fig. 1(c) and (f) display the second PCA component of the raw CSI amplitudes, which display that the CSI correlation can be extracted effectively. Our experimental results show the same CSI correlation extraction results with the existing work (e.g., [13], [14]), when human motion activities occur in the same room with the transmitter and the receiver.

Second, the transmitter and the receiver were deployed in the neighboring rooms, respectively. That is, the wall blocks all the direct and reflected propagation paths between the transmitter and the receiver. Fig. 2 displays the second component of PCA waveforms with three types of human activities (walking, sitting down and falling). Fig. 2(a) displays the PCA waveform of walking, which contains only a part of regular CSI correlation. Fig. 2(b) and (c) display the PCA waveforms of sitting down and falling, respectively. However, the PCA waveforms of sitting down and falling are distorted severely, which cannot extract CSI correlation effectively. Our experimental results show that the existing denoising techniques (low-filter and PCA) could result in arbitrarily bad activity recognition performance, when all the direct and reflected propagation paths between the transmitter and the receiver are blocked by the concrete wall.

##### B. Insight and Motivations

In fact, existing efforts mainly include three steps to distinguish different human activities with CSI information. (1) CSI correlation between human activity and its resulting changes

TABLE I  
RF ATTENUATION IN COMMON BUILDING MATERIALS AT  
2.4 GHz [16]

Building Materials	2.4 GHz
Glass	3 dB
Solid Wood Door 1.75 inches	6 dB
Interior Hollow Wall 6 inches	9 dB
Concrete Wall 18 inches	18 dB

in CSI values is obtained from the raw CSI measurements. (2) Features are extracted from the CSI correlation. (3) The human activity is distinguished by the extracted features. In three steps, the quality of CSI correlation directly determines the accuracy of human activity identification. PCA is considered as a very effective correlation extraction technique, and its correlation decreases when the order of components is increasing [26]. However, existing work selected CSI correlation from the second and subsequent components, and discard the first PCA component (e.g., [13], [14]). Therefore, the two aspects are contradictory to PCA theory. In other words, the CSI correlation extracted from the raw CSI measurements cannot represent the real correlation between human activity and its resulting changes in CSI values. The reason is that the CSI correlation extracted directly from the raw CSI measurements includes not only the correlation of human activity but also the correlation of background environment and noise.

If the human activity occurs in the same room with the transmitter and receiver, the changed CSI values caused by the human activity is relatively strong. This is why CSI correlation can be obtained from PCA waveforms of the raw CSI measurements, but they should choose a useful component from many PCA components. However, the changed CSI values caused by human activity would become very weak, when the wall blocks all the direct and reflected propagation paths between the transmitter and receiver. Therefore, the CSI correlation between human activity and its resulting changes in CSI values would be severely disturbed by the correlation of the background environment and noise. For example, the Wi-Fi signals would be attenuated 18 dB when 2.4 GHz Wi-Fi signals pass through a concrete wall with 18 inches width [16]. Table I displays the signals attenuation by different walls. Meanwhile, our experiment results also display that PCA waveforms are disturbed severely, and CSI correlation cannot be extracted effectively.

To obtain CSI correlation between human activity and its resulting changes in CSI values, can we eliminate the interference of the complicated indoor propagation environment and noise? If we can, the CSI correlation between human activity and its resulting changes in CSI values could be extracted.

### C. CSI Measurement Matrix Analysis

In this subsection, we discuss whether the complicated background environment CSI values can be separated from the following two considerations. (1) Does CSI amplitude matrix of the complicated background environment have low rank characteristics? (2) Is the changed CSI amplitude matrix caused by

human actions sparse? If the rank of background environment CSI matrix is low and the changed CSI matrix caused by human actions sparse, then the complicated background environment CSI component can be separated according to Equation (6).

First, we discuss whether the rank of the complicated background environment CSI matrix is low. We give our experimental results of CSI measurements from different static environments as shown in Fig 3. Fig. 4 displays the CSI amplitudes and correlation matrix of static Environment1. Fig. 4(a), (b) and (c) display the CSI amplitudes of three different subcarriers with different antennas. Experiment results show that the CSI amplitudes of different subcarriers of the same antenna change similarly, whereas different antennas change differently. Fig. 4(d) displays the correlation matrix of CSI amplitudes of 30 subcarriers with the antenna 2. Experiment results show that the correlations of CSI amplitudes of different subcarriers of the same antenna are all greater than 0.5.

Second, we give singular values analysis of CSI amplitude matrix with different time and different environments. Fig. 5 displays the analysis results when both the transmitter and receiver were deployed in the same room. Fig. 5(a) displays the CSI singular values comparison results under different time periods and the same sampling Environment1. Fig. 5(b) displays the singular values comparison results under different sampling environments. We also give the experiment results of CSI singular values analysis when the transmitter and receiver were deployed between the neighboring rooms. Fig. 6(a) displays the CSI singular values comparison results under different time periods and the same sampling Environment1. Fig. 6(b) displays the singular values comparison results under different sampling environments. The experiment results show that the singular values are mainly concentrated on a few elements, whatever the CSI are sampled from different times and different environments. According to our experiment results (Fig. 4, 5, and 6), we can find that the rank of the CSI amplitude matrix of indoor physical environment is approximately low.

Finally, we choose two typical human actions (falling and walking) to verify that the changed CSI matrix caused by human actions is sparse. From the point of the duration of actions, falling is usually a short duration action, while walking has much longer duration. In our experiment, CSI amplitudes were sampled from Environment3 when the transmitter and receiver were deployed in the neighboring rooms with the concrete wall. Fig. 7 displays CSI amplitudes of three different subcarriers on different antennas. We only selected three subcarriers to display because all subcarriers in the same antenna exhibit the similar CSI changes in our experiments. We shall illustrate that the changed CSI matrix caused by human actions is sparse from the following two aspects of our experimental results. (1) The same human action has different influences on different antennas. The subcarriers of some antennas have almost no influence, and some antennas have significant influence. For example, the falling action has little influence on the subcarriers of the first and third antennas, while the subcarriers of the second antenna were influenced severely as shown in Fig. 7(a), (b) and (c). The walking action has a great influence on the subcarriers of the third antenna, and the influences on the first and second

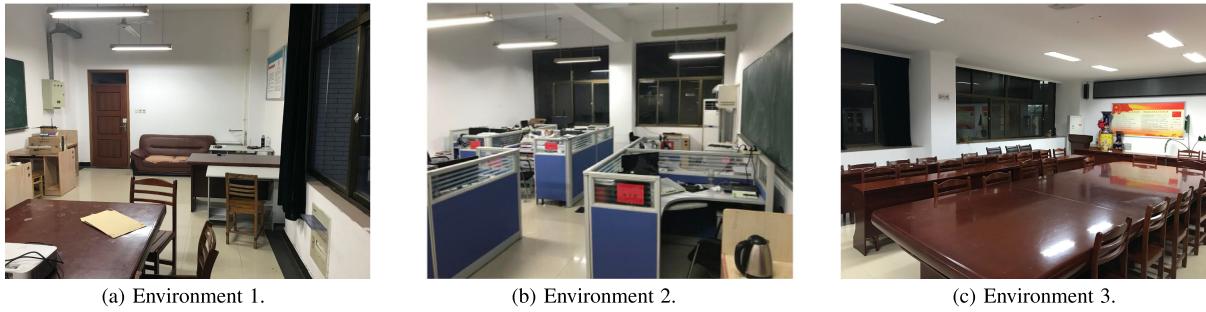


Fig. 3. Three types of CSI data sampling environments.

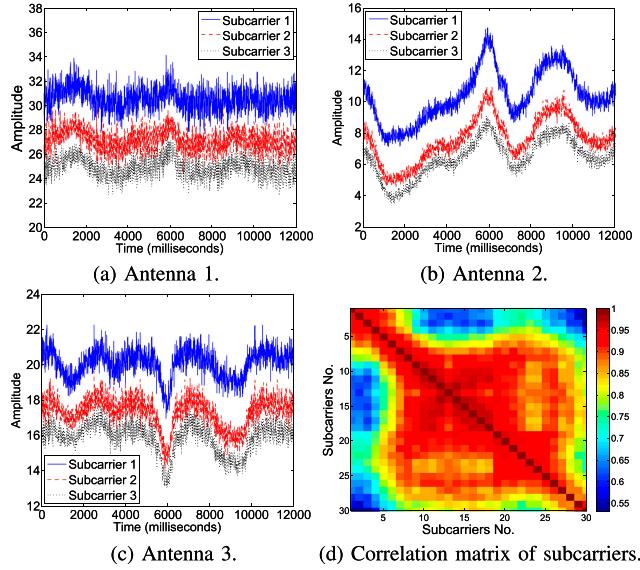


Fig. 4. CSI amplitudes and correlation matrix of static Environment 1.

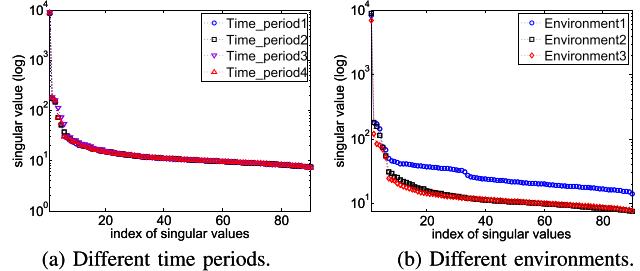


Fig. 5. CSI singular value analysis with the transmitter and receiver in the same room.

antennas are weak as shown in Fig. 7(d), (e) and (f). That is, different human activities only influence a part of subcarriers CSI values, and other parts still retain the characteristics of the background CSI. (2) Although the duration of different human actions is different, the changed CSI time length caused by human action is relatively short if the selected CSI time is long enough. For example, the changed CSI time caused by the falling and walking are different, but the duration time is relatively short compared with our selected CSI window time as shown in Fig. 7(b) and (f). In fact, we can adjust the length of the selected CSI window time according to different human actions.

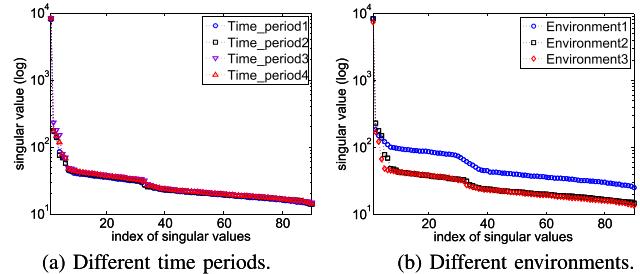


Fig. 6. CSI singular value analysis with the signals through the wall.

However, in general, the duration of indoor human actions is relatively short, and then the selected CSI window time required for activity recognition also need not be too long. Based on the above two considerations, we know that the changed CSI caused by human activity can be considered to be sparse from the dimensions of time and antenna (subcarrier).

According to properties of low rank-matrix decomposition [20], a low-rank matrix with sparse noise addition could be partitioned into a low-rank matrix and a sparse matrix. Therefore, we could divide the raw CSI measurements into the background physical environment CSI values and the changed CSI values by Equation (6).

## V. TW-SEE OVERVIEW

### A. TW-See Description

TW-See consists of two Wi-Fi devices as transmitter and receiver. We use a commercial off-the-shelf Wi-Fi AP device with 1-antenna as the transmitter. We use a common wireless NIC with 3-antennas as the receiver. In our experiments, the transmitter is a Tenda W15E AP, and the receiver is a Mini PC with Intel 5300 NIC. In TW-See, the AP and receiver are deployed in different neighboring rooms. That is, the wall blocks all the direct and reflected propagation paths between the AP and the receiver. If a human activity occurs in either the receiver room or transmitter room, TW-See system can detect it automatically. Fig. 8 displays a brief example of TW-See system.

### B. Main Technical Issues

To realize TW-See system, we need to solve the following technical issues: 1) **Changed CSI values separation and correlation extraction**. Given the input raw CSI measurements, we

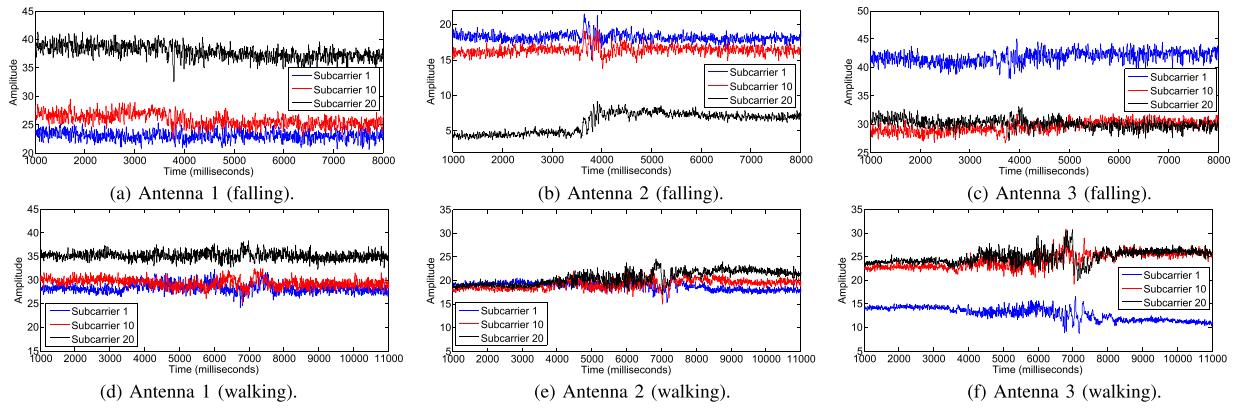


Fig. 7. CSI amplitudes of different subcarriers with falling and walking.

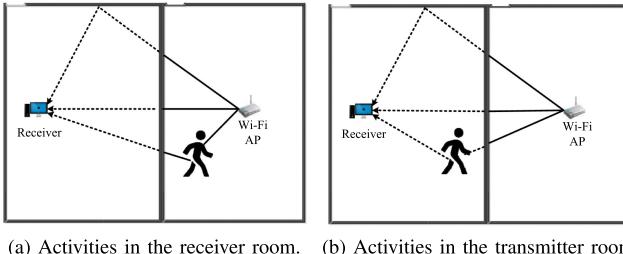


Fig. 8. Brief example of human activity recognition when the signals pass through the wall.

want to separate it into changed CSI values caused by human activity and indoor physical environment CSI values. Meanwhile, we also need to remove the impulse and burst noise of the changed CSI values, and obtain the correlation between human activity and the changed CSI values. 2) **Activity segmentation**. After the correlation extraction, we want to detect the start and end of the activity, and segment it from the CSI correlation waveforms. 3) **Feature extraction and activity recognition**. According to the segmented activity waveforms, we extract activity features and build human activity classifier model.

### C. System Architecture

To solve the aforementioned technical issues, we divide TW-See system into four main blocks as shown in Fig. 9, CSI sampling block, CSI correlation extraction block, activity feature extraction block, and human activity recognition block.

**CSI sampling block:** This block is responsible for sampling CSI measurements from each receiver antenna. If the transmitter with  $N_{tx}$ -antennas and the receiver with  $N_{rx}$ -antennas, then  $N_{tx} \times N_{rx} \times M$  CSI measurements would be sampled in one sample slot, where  $M$  is the number of subcarriers in the OFDM channel between the transmitter and the receiver.

**CSI correlation extraction block:** This block implements three functions. The first function merges CSI measurements of receiver antennas into CSI streams. The second function separates the raw CSI measurements into indoor physical environment CSI values and changed CSI values. The third function

extracts the correlation between the changed CSI values and human activity.

**Activity feature extraction block:** After obtaining the changed CSI correlation waveforms, this block is responsible for detecting the start and end of human activity. Then the pre-defined features would be extracted from the segmented activity waveforms.

**Human activity recognition block:** This block is responsible for distinguishing different human activities. If an activity is recognized successfully, then the activity label will be output. If the recognition result is wrong, the detected activity sample would be added to the training dataset which helps the system to update the classifier model.

## VI. WI-FI SIGNAL PROCESSING

### A. CSI Data Collection

CSI measurements are obtained from received packages in receiver. According to CSI tool introduction in [27], each CSI packet contains a  $N_{tx} \times N_{rx} \times 30$  CSI matrix, where  $N_{tx}$  and  $N_{rx}$  are the number of antennas of the transmitter and the receiver, respectively. 30 is the number of subcarriers in the OFDM channel between the transmitter and the receiver. In TW-See, the transmitter has 1-antenna and the receiver has 3-antennas. Therefore, each packet contains  $1 \times 3 \times 30$  CSI measurements. To simplify CSI data process, CSI measurements of the  $t$ -th packet are reformed as  $CSI_t = [CSI_t^1, CSI_t^2, CSI_t^3]^T$ , where  $CSI_t^i$  is the CSI measurements of the  $i$ -th antenna of receiver.  $CSI_t^i = [csi^{(i,1)}, csi^{(i,2)}, \dots, csi^{(i,30)}]^T$ , where  $CSI_t^{(i,j)}$  is the CSI measurement of the  $i$ -th antenna and  $j$ -th subcarrier. Therefore,  $N$  CSI packets could be denoted as  $CSI\_streams = [CSI_1, CSI_2, \dots, CSI_N]_{90 \times N}$ .

In TW-See, the CSI sampling rate is 1000 packets/s, namely, we can collect 1000 CSI measurements for each 90 CSI streams in one second. Since the CSI phase is affected by carrier frequency offset [28], we only consider the CSI amplitude in TW-See system.

### B. Low-Pass Filter Denoising

In TW-See, the CSI measurements obtained from commercial Wi-Fi NIC, which contain noise from various sources

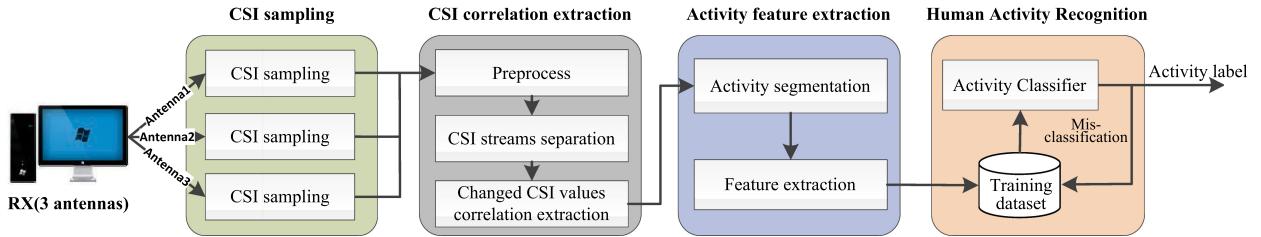


Fig. 9. Architecture of TW-See system.

such as interference coming from nearby devices, transmission power adaptation at the transmitter, and imperfect clock synchronization [14]. Therefore, we should remove the noise from the raw CSI measurements first. Considering that CSI waveforms obey the principle of superposition of waves, CSI frequency is determined by the moving speed of human motion, namely,  $f = 2\nu/\lambda$ , where  $f$  is the frequency of the CSI waveform,  $\nu$  is the moving speed, and  $\lambda$  is the wavelength of radio signals. For example, a moving speed of 1.5 m/s can be converted to the frequency of 51.8 Hz when the radio wavelength  $\lambda$  is 5.79 cm. So, we choose the Butterworth low-pass filter as our CSI denoising technique. In our experiment, we set the band threshold of low-pass filter to 100 Hz.

## VII. OR-PCA-BASED CSI CORRELATION EXTRACTION

### A. Changed CSI Value Separation

According to our experiment results, CSI correlation between human activity and its resulting changes in CSI values cannot be obtained from the raw CSI measurements by PCA. There are mainly the following three reasons. (1) The changed of CSI values caused by human activity become very slight after being blocked by the wall, which causes them to be susceptible to noise. (2) The complicated indoor propagation environment seriously interfered with the distribution of the changed CSI values. (3) The original PCA technique is severely affected by noise during data correlation extraction.

According to our CSI amplitudes matrix analysis and low-rank matrix decomposition theory, we can partition the raw CSI measurements into two components, the indoor physical environment CSI values and the changed CSI values. In mathematics, the process of separation can be denoted as

$$\begin{aligned} \min \quad & \|CSI_{\text{bg}}\|_* + \lambda \|CSI_{\text{act}}\|_1 \\ \text{s.t. } & CSI_{\text{raw}} = CSI_{\text{bg}} + CSI_{\text{act}} \end{aligned} \quad (7)$$

where  $CSI_{\text{raw}}$  is the raw CSI measurements,  $CSI_{\text{bg}}$  and  $CSI_{\text{act}}$  are the background physical environment CSI values and the changed CSI values, respectively. In Equation (6),  $E$  is equivalent to  $CSI_{\text{act}}$ , which is the sparse noise and should be discarded. However,  $CSI_{\text{act}}$  contains the changed CSI values caused by human activity and noise. To solve the problem (7), we choose the augmented lagrange multiplier method [24] as our algorithm.

Fig. 10 displays raw CSI measurements and its low rank separation results. In our experiments, the parameter  $\lambda$  is set to 0.1. Because the CSI amplitudes of different subcarriers in

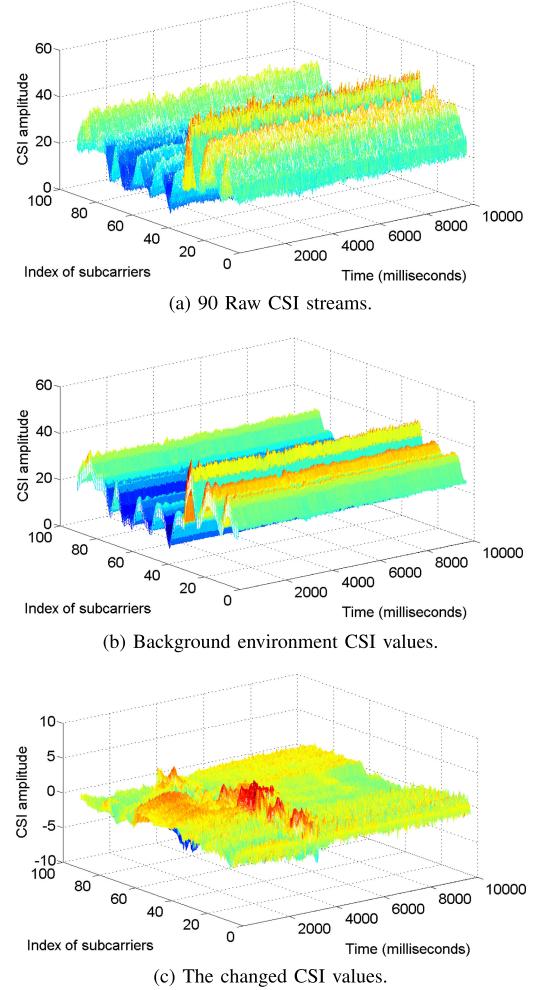


Fig. 10. 90 Raw CSI streams and low rank separation results.

the same antenna have the strong correlation and the receiver has three antennas, we set the rank of the background CSI matrix to 3. In our collected CSI data, the rank of the separated CSI background matrix is nearly 3 when  $\lambda$  is set to 0.1. Fig. 10(a) displays the raw CSI measurements corresponding to the fall activity, which is difficult to find any regular change. Fig. 10(b) and (c) show the background physical environment CSI signals and the changed CSI signals, respectively. We can find obvious fluctuation of CSI amplitudes easily as shown in Fig. 10(c).

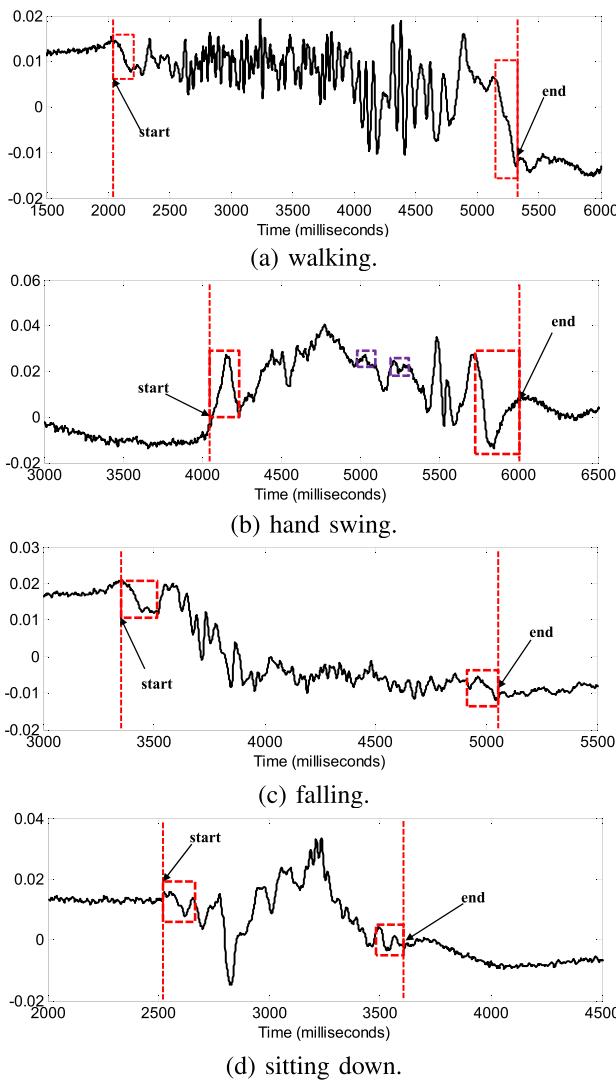


Fig. 11. The first component of Or-PCA waveforms with different activities when the signals pass through the wall.

### B. CSI Correlation Extraction

After separating the raw CSI measurements, we can obtain the background CSI values and the changed CSI values. According to low-rank matrix decomposition theory, the changed CSI values should be discarded as noise or abnormal values. In TW-See, however, the CSI information of human activity is included in the changed CSI values, the CSI correlation should be extracted from the changed CSI values. Therefore, the original low-rank matrix decomposition cannot be applied directly to our CSI correlation extraction, and we call our proposed approach as the opposite robust PCA (Or-PCA).

Indeed, we only want to extract the CSI correlation between human activity and its resulting changes in CSI values from  $CSI_{act}$ . Then we exploit Or-PCA to extract the CSI correlation from  $CSI_{act}$ . Fig. 11 displays the first component of Or-PCA waveform with four types of activities, when the transmitter and receiver of Wi-Fi signals were deployed in the different rooms. Our experimental results show that the Or-PCA wave-

form changes significantly when human motion activity occurred, even if the Wi-Fi signals pass through the wall. There is a clear distinction between the waveforms which exploits PCA on the raw CSI measurements as shown in Fig. 2.

Compared with the existing work, another advantage of our approach is that the CSI correlation is mainly concentrated on the first PCA component. In the existing work, the first PCA component is discarded because it cannot extract the useful CSI correlation. The reason is that the first PCA component mixes the principal component of indoor physical environment CSI values, when PCA is used on the raw CSI measurements directly. Meanwhile, the indoor physical environment CSI values and sparse noise disturb the distribution of the principal component of the CSI correlation. This is why the CSI correlation between human activity and its resulting changes in CSI values is distributed on the first 5-20 principal components, when the PCA is directly applied to the raw CSI data.

## VIII. HUMAN ACTIVITY RECOGNITION

### A. Activity Segmentation

Although we exploit Or-PCA to obtain the CSI correlation between human activity and its resulting changes in CSI values, it is still difficult to detect the end of the activity. Since the wall blocks Wi-Fi signals power severely, the Or-PCA waveform does not change dramatically during the human activity time. In our experiments, we test six types of activities signals (walking, hand swing, sitting down, falling, boxing and standing up) through the wall. Fig. 11 displays CSI waveforms after robust Or-PCA. Fig. 11(a) display that the CSI waveforms of walking changes dramatically during the activity time. Fig 11(b), (c) and (d) display that the CSI waveform of hand swing, sitting down and falling does not always change dramatically during the activities time. For example, the CSI waveforms of the purple blocks change smoothly and flatly as shown in Fig 11(b). Although existing techniques, such as exponential moving average algorithm (e.g., [14]) and threshold-based sliding window algorithm (e.g., [15]), could segment the activity when its Or-PCA amplitudes change dramatically, these techniques are difficult to detect the end of activity when the Or-PCA amplitudes change smoothly and flatly during human activity.

To localize the start and the end of activities effectively, we propose a normalized variance sliding windows algorithm as shown in Algorithm 1. The idea of Algorithm 1 mainly includes two aspects. (1) We normalize the variances of Or-PCA waveforms, which makes it possible to use the same threshold value for activity time detection, and eliminate the influence of different amplitudes of the Or-PCA waveforms on the human activity segmentation. (2) We exploit time sliding window to distinguish the true or false of the end activity time. The input parameters of Algorithm 1 are  $g(t)$ ,  $\delta$ ,  $W_t$ ,  $\phi$ , and  $\mu$ .  $g(t)$  is the activity waveforms after Or-PCA based correlation extraction.  $\delta$  is a threshold value, which is used to detect the start and the end of the activity.  $W_t$  is the size of time sliding window.  $\phi$  and  $\mu$  are the weighted parameters, which are used to determine whether the human action is over, not the intermediate state of the action. The output is the start time  $t_{start}$  and

**Algorithm 1:** Activity Segmentation Algorithm.

---

```

Input : Or-PCA waveform  $g(t)$ ,  $\delta$ ,  $W_t$ ,  $\phi$ ,  $\mu$ 
Output: activity start time  $t_{\text{start}}$  and end time  $t_{\text{end}}$ 

1  $i \leftarrow 1;$ 
2 while  $t(i * W_t)$  not end do
3    $V(i) \leftarrow \text{Var}(g(i * W_t) - g((i - 1) * W_t));$ 
4    $i \leftarrow i + 1;$ 
5    $V_n \leftarrow \text{Normalized}(V);$ 
6    $flag\_start \leftarrow \text{false};$ 
7    $i \leftarrow 1;$ 
8   for  $i \leq \text{length}(V_n)$  do
9     if  $flag\_start == \text{false}$  then
10       if  $V_n(i) \geq \delta$  then
11          $t_{\text{start}} \leftarrow (i - 1) * W_t;$ 
12          $flag\_start \leftarrow \text{true};$ 
13     else
14       if  $V_n(i) < \delta$  then
15          $V_{\text{test}} \leftarrow \phi * V_n(i) + (1 - \phi) * V_n(i + 1);$ 
16         if  $V_{\text{test}} < (\mu * V_n(i + 1))$  then
17            $t_{\text{end}} \leftarrow i * W_t;$ 
18           break;
19      $i \leftarrow i + 1;$ 
20 return  $t_{\text{start}}, t_{\text{end}};$ 

```

---

the end time  $t_{\text{end}}$  of the detected activity. In Algorithm 1, line 2–4 is used to calculate the variances of the Or-PCA waveforms at different times.  $\text{Var}(g(i * W_t) - g((i - 1) * W_t))$  calculates the variance of Or-PCA waveforms in the  $i$ -th time windows. In our experiments,  $W_t$  is set to 200 ms. Line 5 normalizes the Or-PCA waveform variances. Line 10–12 is used to determine whether the human action starts or not. Line 14–18 is used to determine whether the human action is over or not.  $V_{\text{test}} = \phi * V_n(i) + (1 - \phi) * V_n(i + 1)$  is used to calculate the weighted value of variance between the two consecutive time windows, which determines whether the end of the activity is true or not. According to our experiment results, we set  $\delta = 0.2$ ,  $\phi = 0.9$ , and  $\mu = 3$ .

**B. Feature Extraction**

According to Algorithm 1, we could segment the human activity from Or-PCA waveform. Next, we would like to extract the concreted features from the segmented activity Or-PCA waveform. In TW-See, we choose the following eight features as candidate features. (1) the normalized STandard Deviation of CSI (STD), (2) the Median Absolute Deviation (MAD), (3) Interquartile Range (IR), (4) the Velocity of Signals Change (VSC), (5) Duration of Activity (DA), (6) the MIn value of PCA Amplitude (MIA), (7) the MAx value of PCA Amplitude (MAA), and (8) the MEan value of PCA Amplitude (MEA).

**C. BP Neural Network**

In TW-See, we exploit Back Propagation Neural Network (BP Neural Network) [29] to recognize human activities among

the segmented activities by the features extracted above. Compared to other classification techniques, neural networks not only learn automatically from data sets but also compute any function [30]. Therefore, BP Neural Network is very suitable for TW-See system, which needs to solve multi-activity classification problem. To implement a neural network that learns to recognize human activities, we need to build a training dataset and test dataset. In the training dataset building phase, human activities (walking, sitting down, falling, hand swing, boxing and standing up) CSI signals are sampled from a given training indoor environment. Our training dataset does not change even if the detection environment is changed. After sampling the training CSI measurements, the aforementioned eight features are extracted from the Or-PCA waveform of the segmented human activity. Meanwhile, each human activity sample needs to be labeled, then using the training dataset and their labels to train our BP Neural Network. Consider this is a multi-classification problem, the One-Hot Encoding method is applied in the label phase because it fits encode discrete categories.

According to the extracted features, we design a three-layer BP Neural Network in our system, including an input layer with 24 neurons, an output layer with 7 neurons and a hidden layer with 14 neurons. In our BP Neural Network, the sigmoid function is selected as the activation function, and the cost function is the cross-entropy cost function. We initialize the weighted values as Gaussian random variables with a mean of 0 and a standard deviation of  $1/\sqrt{n_{in}}$ , where  $n_{in}$  is the number of input weights of a neuron. Meanwhile, the bias is initialized as Gaussian random variables with a mean of 0 and a standard deviation of 1, the learning rate is 0.03, the mini-batch size is 10, the number of training epochs is 110.

**IX. IMPLEMENTATION AND EVALUATION****A. Implementation**

TW-See system includes a laptop or PC with a WiFi NIC and a commercial Wi-Fi AP. We implemented TW-See system using a mini PC with the Ubuntu system, the Wi-Fi NIC is the Intel Wi-Fi Link 5300 with 3 antennas as the receiver, and the AP is the Tenda W15E Router with one antenna as the transmitter. To collect the CSI measurements, we installed the Linux CSI tool [27] on our PC. In our experiments, the transmitter was set to work at 5 GHz band and used the 165-th channel of 20 MHz channel bandwidth. Meanwhile, TW-See acquires CSI measurements from the Linux CSI tool and processes it in real-time using MATLAB. In TW-See, the CSI measurements sampling rate is set to 1000 packets/s.

**B. Experimental Setup**

We evaluate the performance of TW-See in three different environments as shown in Fig. 12 and 13. The walls of environments 1 and 2 are the concrete wall with 12 inches. In environment 3, the top of the wall is glass, and the bottom is concrete as shown in Fig. 13. The experimental environment 1 is a meeting room and an office room which are neighboring to each other. The meeting room is 12 m in length and 8.7 m

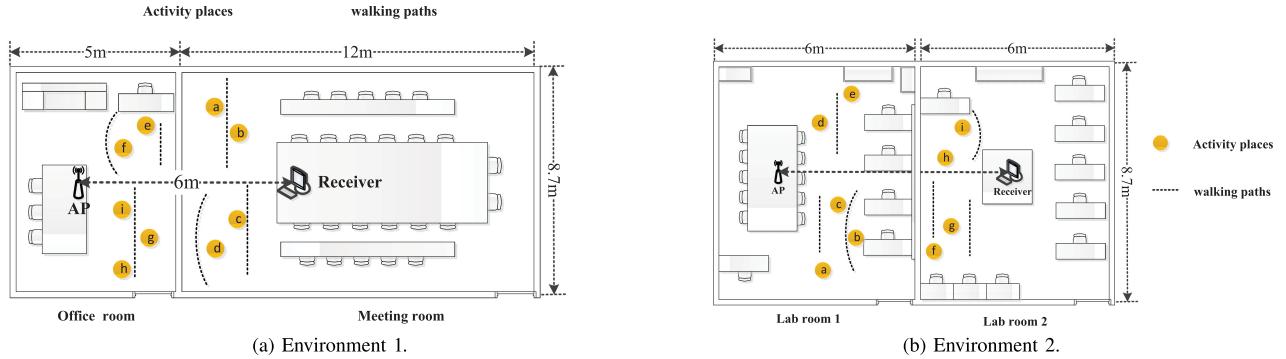


Fig. 12. Floor plans of two experimental environments. (a) Environment 1 (a meeting room and an office room). (b) Environment 2 (two lab rooms).

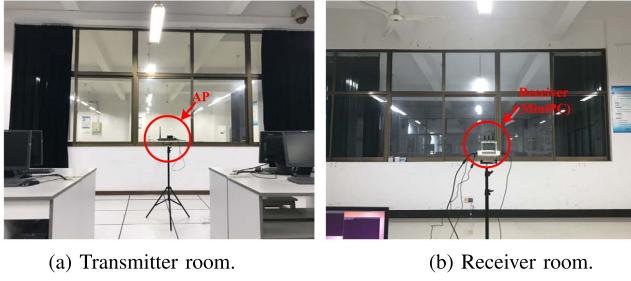


Fig. 13. Experimental environment 3 with glass wall.

in width. The office room is 5 m in length and 8.7 m in width. The experimental environment 2 is two lab rooms which are neighboring to each other. The two lab rooms are 6 m in length and 8.7 m in width. The experimental environment 3 is a lab room and an empty room which are neighboring to each other. The lab room and empty room are 6 m in length and 8.7 m in width.

In environment 1, the PC with Intel Wi-Fi Link 5300 was deployed in the meeting room, and Tenda W15E Router was deployed in the neighboring office room. In environment 3, the PC with Intel Wi-Fi Link 5300 was deployed in the lab room, and Tenda W15E Router was deployed in the neighboring empty room. The distance of three environments between the transmitter and receiver is 6 m. The environment 1 is mainly used to evaluate the accuracy of TW-See system. The environment 2 and 3 are used to evaluate the robustness of our system. We chose the test environment 2 and 3 that are entirely different from the training rooms. The test environment 2 are two lab rooms with the concrete wall, which contain many workbenches and the wall covered with blackboards as shown in Fig. 12(b). The test environment 3 are two neighboring rooms with the glass wall. In our experiments, the training and testing samples are shown in Table II.

### C. $\lambda$ Parameter

According to Equation (7), the raw CSI measurements is divided into two components, background CSI matrix and the changed CSI matrix caused by human action. The division of two components is determined by the parameter  $\lambda$ . Meanwhile,

TABLE II  
THE TRAINING SAMPLES AND TEST SAMPLES

Human activity	Training samples	Test samples			Label
		Env. 1	Env. 2	Env. 3	
walking	240	80	80	80	1000000
sitting down	240	80	80	80	0100000
standing up	240	80	80	80	0010000
falling	240	80	80	80	0001000
hand swing	240	80	80	80	0000100
boxing	240	80	80	80	0000010
empty	240	80	80	80	0000001

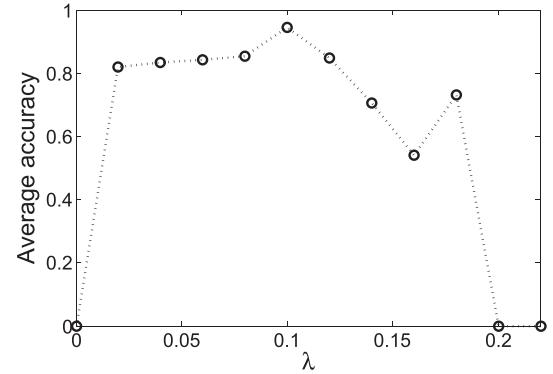


Fig. 14. Average accuracy of different human activities with different  $\lambda$  in environment 1.

the correlation between human activity and its resulting changes in CSI values is only obtained from the separated changed CSI matrix. This correlation directly determines the recognition accuracy of human activity. Therefore, the value of  $\lambda$  affects the performance of TW-See. In our experiments, we calculated the average accuracy of seven types of human activities in environment 1. Every human activity contains 240 training samples and 80 testing samples as shown in Table II. The value of  $\lambda$  was set from 0 to 0.2 with the step 0.02.

Fig. 14 displays the average accuracy of different human activities with different values of  $\lambda$ . Our experimental results show that the average accuracy is the highest (94.46%) when the value of  $\lambda$  is near 0.1. We find the raw CSI measurements can't be separated when the  $\lambda$  is set to greater than or equal to 0.2. The average accuracies are nearly above 80% when

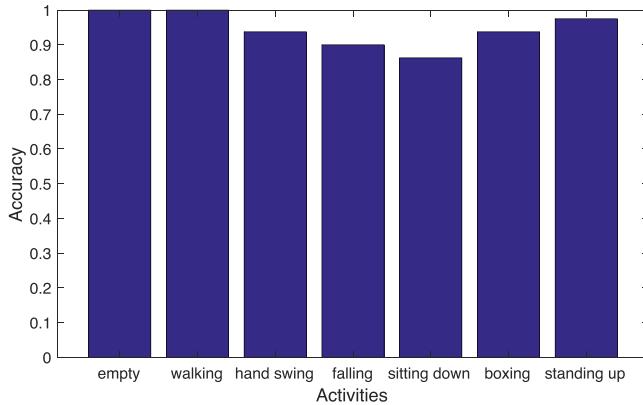


Fig. 15. Experiment results of recognition accuracy with different human activities in environment 1.

the value of  $\lambda$  is set between 0.02 and 0.12. Our experimental result is also the same as our analysis of the parameter  $\lambda$  (Section VII-A). The rank of the separated background CSI matrix is 3 when the value of  $\lambda$  is near 0.1 in our experiment environment. Therefore, the value of  $\lambda$  was set to 0.1 in the following our experiments.

#### D. Accuracy

In this subsection, we evaluate the average cross-validation accuracy of TW-See system when the transmitter and receiver were deployed between a meeting room and an office room (environment 1). In our experiments, we recruited 6 volunteers with 4 males and 2 females to participate in the CSI data collection. We collected seven kinds of human activities, which are walking, sitting down, standing up, falling, hand swing, boxing and empty (i.e., no activity). These activities include both large movements such as walking and small movements such as hand swing. We collected 320 samples for each activity, and the training dataset contains 1680 samples, and the validation dataset contains 560 samples. After obtaining the training and validation dataset, we extracted the activity features and trained our BP Neural Network based on our proposed approaches.

Fig. 15 displays the human activity recognition accuracy of 4-fold cross-validation in the meeting room and office room. TW-See achieves an average cross-validation accuracy of 94.46% across all seven human activities. The accuracy of empty, walking, hand swing, falling, sitting down, boxing and standing up is 100%, 100%, 93.75%, 90%, 86.25%, 93.75%, and 97.5%, respectively. In our system, the accuracy of walking and empty is 100%. In other words, walking and empty cannot be recognized as the other human activities. Meanwhile, the other human activities also cannot be recognized into the walking or empty. The experimental results show that TW-See can achieve high recognition accuracy for different human activities.

#### E. Robustness

Since the wireless signals are susceptible to the environment, we evaluate the robustness of TW-See system from three

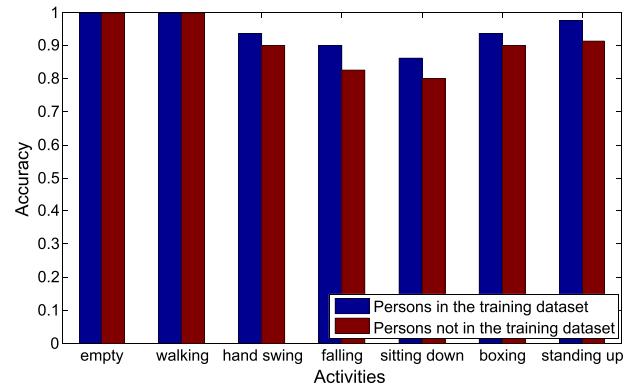


Fig. 16. Experiment results of different testing persons.

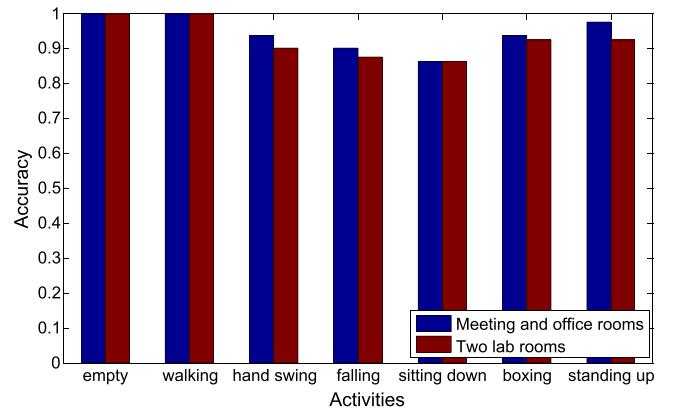


Fig. 17. Experiment results of concrete wall with different environments.

aspects: (1) different training and testing participation persons in environment 1, (2) different experiment environments with the same wall material (environment 2); (3) different experiment environments with different wall material (environment 3).

First, we evaluate the robustness of our system when the volunteers of the testing dataset are not in the training dataset. We collected 80 samples for each activity by three volunteers in the training dataset, where the three volunteers were not in the training dataset. In this case, TW-See achieves an average accuracy of 90.54%. Fig. 16 displays the comparison results of accuracy with persons in the training dataset and persons not in the training dataset. When the persons involved in the testing were different with the training, the accuracy of walking and empty is also both 100%, the accuracy of hand swing is 90%, the accuracy of falling is 82.5%, the accuracy of sitting down is 80%, the accuracy of boxing is 90%, and the accuracy of standing up is 91.25%. Our experimental results show that TW-See system has good activity recognition performance whether the test environment or the participants change or not.

Second, we evaluate the robustness of TW-See without changing our training dataset when the testing samples were collected from two lab rooms (environment 2). In environment 2, we collected 80 samples for every activity. TW-See system achieves an average accuracy of 92.68%. Fig. 17 shows the comparison results of accuracy with different environments for each activity. In two lab rooms, the accuracy of walking and empty is both

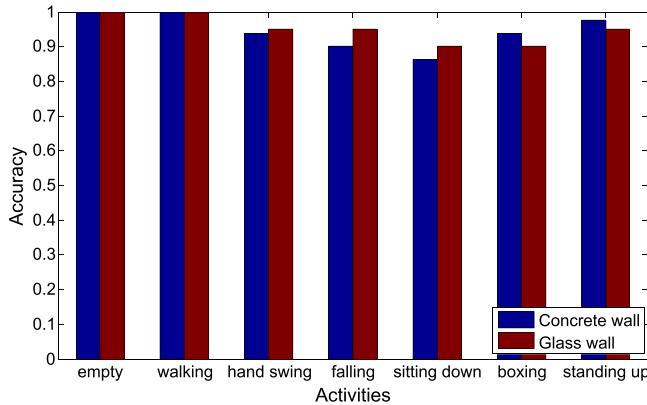


Fig. 18. Experiment results of different material walls.

100%, the accuracy of hand swing is 90%, the accuracy of fall is 87.5%, the accuracy of sitting down is 86.25%, the accuracy of boxing is 92.5%, and the accuracy of standing up is 92.5%. Although the indoor experiment environment is very different from the test environment, our experimental results show that the performance of human activity recognition is still relatively accurate.

Finally, we evaluate the robustness of TW-See with different wall materials. In environment 3, we collected 320 samples for every activity, of which 240 samples were used for training, and 80 samples were used for testing. Because different material walls have different signal attenuation, we exploited samples of environment 3 as the training dataset. However, the BP Neural Network model and hyperparameters are the same as environment 1. In environment 3, TW-See achieves the average accuracy is 95%, the accuracy of seven activities is 100%, 100%, 95%, 95%, 90%, 90% and 95%, respectively. According to our experimental results as shown in Fig. 18, the recognition accuracy of handing swing, falling and sitting down in the concrete wall is less than that in glass wall. Meanwhile, the recognition accuracy of boxing and standing up in the concrete wall is better than that in the glass wall. The main reasons include the following two aspects. One aspect is that the signal attenuation by the glass wall is less than that of the concrete wall. Another aspect is that we exploited the same classification model and hyperparameters without considering whether the environment has changed. Therefore, the classification model and hyperparameters are not optimal for the environment 3.

According to the above experimental results, TW-See can achieve high activity recognition accuracy and good robustness performance for different environments.

## X. CONCLUSION

In this paper, we proposed TW-See, a device-free passive human activity recognition system with Wi-Fi signals. TW-See could meet the requirement of the received signals through the wall. To eliminate the inference of the complicated indoor environment on CSI correlation extraction, we proposed an Or-PCA approach to extract the correlation between human activity and

its resulting changes in CSI values. Unlike existing approaches, Or-PCA can not only effectively extract the correlation but also make it concentrated on the first principal component. According to the obtained correlation, we proposed a normalized variance sliding windows algorithm to segment Or-PCA waveform. After Or-PCA waveform segmentation, we extracted the activity features and use a BP neural network to recognize different human activities. Meanwhile, We implemented TW-See using commercial WiFi devices and evaluated it in different environments. Our results show that TW-See achieved average 94.46% accuracy when the signals pass through the wall.

## REFERENCES

- [1] K. Yatani and K. N. Truong, "Bodyscope: A wearable acoustic sensor for activity recognition," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 341–350.
- [2] D. Fortin-Simard, J. S. Bilodeau, K. Bouchard, S. Gaboury, B. Bouchard, and A. Bouzouane, "Exploiting passive RFID technology for activity recognition in smart homes," *IEEE Intell. Syst.*, vol. 30, no. 4, pp. 7–15, Jul./Aug. 2015.
- [3] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Comput. Sci.*, vol. 34, pp. 450–457, 2014.
- [4] G. Debard *et al.*, "Camera-based fall detection on real world data," in *Outdoor and Large-Scale Real-World Scene Analysis* (Lecture Notes in Computer Science, vol. 7474), F. Dellaert, J.-M. Frahm, M. Pollefeys, L. Leal-Taix, B. Rosenhahn, Eds. Berlin, Germany: Springer, 2012, pp. 356–375.
- [5] H. Foroughi, B. S. Aski, and H. Pourreza, "Intelligent video surveillance for monitoring fall detection of elderly in home environments," in *Proc. Int. Conf. Comput. Inf. Technol.*, 2008, pp. 219–224.
- [6] A. Jalal, Y. H. Kim, Y. J. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern Recognit.*, vol. 61, pp. 295–308, 2016.
- [7] Q. Wu, Y. D. Zhang, W. Tao, and M. G. Amin, "Radar-based fall detection based on Doppler time-frequency signatures for assisted living," *IET Radar Sonar Navigat.*, vol. 9, no. 2, pp. 164–172, Feb. 2015.
- [8] S. Tao, M. Kudo, and H. Nonaka, "Privacy-preserved behavior analysis and fall detection by an infrared ceiling sensor network," *Sensors*, vol. 12, no. 12, pp. 16920–16936, 2012.
- [9] Y. Li, K. C. Ho, and M. Popescu, "A microphone array system for automatic fall detection," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1291–1301, May 2012.
- [10] F. Adib and D. Katabi, "See through walls with WiFi," in *Proc. ACM SIGCOMM Conf. SIGCOMM*, 2013, pp. 75–86.
- [11] Q. Pu, S. Jiang, and S. Gollakota, "Whole-home gesture recognition using wireless signals," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 485–486, 2013.
- [12] 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. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.
- [13] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.
- [14] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 363–373.
- [15] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 511–526, 2017.
- [16] "How signal is affected," City of Cumberland Rep., MD, USA. [Online]. Available: [www.cumberlandmd.gov/DocumentCenter/View/962/Signal-Information-PDF?bidId=](http://www.cumberlandmd.gov/DocumentCenter/View/962/Signal-Information-PDF?bidId=)
- [17] A. M. Khan, Y. K. Lee, S. Y. Lee, and T. S. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in *Proc. Int. Conf. Future Inf. Technol.*, 2010, pp. 1–6.
- [18] T. S. Rappaport, "Wireless Communications: Principles and Practice," Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

- [19] D. Zhang, J. Ma, Q. Chen, and L. M. Ni, "An RF-based system for tracking transceiver-free objects," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2007, pp. 135–144.
- [20] J. Wright, A. Ganesh, S. Rao, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices," *Adv. Neural Inf. Process. Syst.*, vol. 87, no. 4, pp. 2080–2088, 2009.
- [21] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Sparse and low-rank matrix decompositions," in *Proc. Allerton Conf. Commun., Control, Comput.*, 2009, pp. 962–967.
- [22] M. Fornasier, H. Rauhut, and R. Ward, "Low-rank matrix recovery via iteratively reweighted least squares minimization," *SIAM J. Optim.*, vol. 21, no. 4, pp. 1614–1640, 2010.
- [23] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, no. 1/2, pp. 321–353, 2011.
- [24] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," unpublished paper, 2010. [Online]. Available: <https://arxiv.org/abs/1009.5055v1>
- [25] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [26] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, no. 1, pp. 37–52, 1987.
- [27] 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.
- [28] J. Gjengset, J. Xiong, G. McPhillips, and K. Jamieson, "Phaser: Enabling phased array signal processing on commodity WiFi access points," in *Proc. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 153–164.
- [29] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [30] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.



**Xiangou Wu** received the Ph.D. degree from the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China, in 2013. He is currently an Associate Professor with the School of Computer Science and Technology, University of Anhui Technology, Maanshan, China. His research interests include wireless sensor networks, crowdsourcing, privacy protection, and datacenter networks.



**Zhaobin Chu** received the B.S. degree in computer science and technology from the University of Anhui Technology, China, in 2016. He is currently working toward the master's degree at the School of Computer Science and Technology, University of Anhui Technology, Maanshan, China. His research interests include wireless sensing and machine learning.



**Panlong Yang** received the B.S., M.S., and Ph.D. degrees in communication and information system from the Nanjing Institute of Communication Engineering, China, in 1999, 2002, and 2005, respectively. He is currently a Professor with the University of Science and Technology of China, Anhui, China. His research interests include wireless mesh networks, wireless sensor networks, and cognitive radio networks. He is a member of the IEEE Computer Society and the ACM SIGMOBILE Society.



**Chaocan Xiang** received the B.S. and Ph.D. degrees in computer science and engineering from the Institute of Communication Engineering, PLA University of Science and Technology, Jiangsu, China, in 2009 and 2014, respectively. He is currently an Assistant Professor with the Army Logistics University of PLA, Chongqing, China. His current research interests include mobile computing, crowd-sensing networks, and IoT.



**Xiao Zheng** received the B.S. degree from Anhui University, Anhui, China, in 1997, and the Ph.D. degree in computer science and technology from Southeast University, Jiangsu, China, in 2014. He is currently a Professor with the School of Computer Science and Technology, Anhui University of Technology, Anhui, China. His research interests include service computing and mobile cloud computing.



**Wenchao Huang** received the B.S. and Ph.D. degrees in computer science from the University of Science and Technology of China, Anhui, China, in 2006 and 2011, respectively. He is currently an Associate Professor with the School of Computer Science and Technology, University of Science and Technology of China. His current research interests include information security, trusted computing, formal methods, mobile computing, and crowdsourcing.