

# SEARE: A System for Exercise Activity Recognition and Quality Evaluation Based on Green Sensing

Fu Xiao, Member, IEEE, Jing Chen, Xiaohui Xie, Linqing Gui, Lijuan Sun, and Ruchuan Wang

**Abstract**—Green-computing technology and energy-saving design have become the focus of research in various fields in recent years. As a ubiquitously deployed infrastructure, WiFi can be considered as a platform for green sensing, and a plethora of efforts have been made in WiFi-based passive detection recently. However, little work has been done on the exercise activity recognition. In this paper, we propose SEARE, a novel energy-efficient solution using WiFi for exercise activity recognition. It is prototyped by fine-grained CSI extracted from existing commercial WiFi devices. Different from traditional features like mean or max value exploited in previous activity recognition works, involving either time or frequency information, we select CSI-waveform shape as activity feature, which contains the information from both of these two domains. A series of de-noise methods are designed, including low-pass, PCA, and median filtering, where PCA can remove the in-band noise that traditional low-pass filters fail to do. Finally the evaluation of activities quality can be made. Extensive experimental result validates the great performance of SEARE in both LOS and NLOS scenarios, with average recognition accuracies of 97.8% and 91.2% respectively.

**Index Terms**—Green Sensing, WiFi, Channel State Information (CSI), Exercise Activity Recognition, Quality Evaluation.

## 1 INTRODUCTION

Recently, with the attention to sustainable development increasing, all aspects about green concept namely energy conservation have become the focus of academic and industrial researchers, turning green computing a new round of hot field. The advent of green computing provides existing researches with brand-new ideas. Thus a number of green technologies are under study, such as green energy management for systems, energy-saving design for personal devices and so on, which will greatly change the way we investigate.

In wireless sensor networks, sensing is the basis of many applications. As for traditional sensor networks, extra dedicated equipment are required for sensing, which is certainly not a green-sensing solution. With the popularity of WiFi devices and the wide deployment of WiFi networks, ubiquitous WiFi can be considered as a good platform for green sensing [1]. Lately, the advancement in WiFi techniques has spurred a

range of related applications, including localization [2], human detection, gesture recognition, etc. Activity recognition as one key technique for many applications is widely used in family medical care for the elderly and human-computer interaction. Its main principle is that human activities cause signals changes in the time series, so we can demodulate the activities information by analyzing modulated received signals. Activity recognition methods primarily fall into two categories: device-based and device-free.

Traditional activity recognition methods are mainly device-based, such as audio or radio based, vision based and sensor based [3]. Audio or radio based methods [4], [5] depend on significant changes of signals, so they are not appropriate for micro-activities recognition. Vision based methods [6], like Kinect [7], highly depend on external environment, such as adequate light and line of sight condition. Sensor based methods [8]- [10] require users to carry specific devices, namely RFID [11], [12], gloves, or mobile sensors, which might be inconvenient in some scenes, for instance, swimming and bathing, and also increase cost.

Currently, researchers have turned to investigate device-free methods based on wireless signals to overcome the problems mentioned above. WiFi as a kind of ubiquitous wireless infrastructure attracts researchers' interest, and a plethora of activity recognition techniques are proposed, such as WiSee [13], WiHear [14], WiKey [15], and Wi-Vi [16]. WiSee is able to recognize nine gestures using Universal Software Radio Peripheral (USRP) that can obtain Orthogonal Frequency Division Multiplexing (OFDM) signals and extract Doppler shift due to the movement of body or limbs. WiHear and WiKey are usually applied in specific scenes. Wi-Vi utilizes virtual multi-antenna technology to detect signal fluctuations contributed from human activities. Gesture recognition [17] and breath detection [18]- [21] have made great progress in recognition accuracy and granularity, and [22] shows moving human detection even in through-wall scene can also be of

• F. Xiao, L. Sun, and R. Wang are with the Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Key Lab of Broadband Wireless Communication and Sensor Network Technology of Ministry of Education and College of Computers, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: xiaof@njupt.edu.cn; sunlj@njupt.edu.cn; wangrc@njupt.edu.cn).

• L. Q. Gui is with the Department of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, 210094 China (e-mail: guilinqing@gmail.com).

• J. Chen and X.H. Xie are with the College of Computers, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: m18020129701@163.com; xiexh\_njupt@163.com).

Manuscript received August 31, 2017; revised November 28, 2017. This work is supported by National Science Foundation of China under grants No. 61373137, 61572261, 61572260; Nature Science Foundation of Jiangsu for Distinguished Young Scientist under Grant BK20170039; The Research Innovation Program for College Graduates of Jiangsu Province under grant No. KYLX16\_0666, KYLX16\_0670 and Scientific and Technological Support Project of Jiangsu Province under grant No. BE2014718, BE2016777.

high accuracy.

Although previous studies have made great breakthrough, few effort has been made on exercise activity recognition, a noteworthy field. Exercise has been an indispensable part of daily life nowadays. Standard movements are conducive to good physical fitness, while irregular ones may be harmful. Motivated by research gap in the field of exercise activities recognition and enlightened by green concept, our work turns to exercise activities recognition based on green sensing.

In the paper, we propose SEARE, a CSI-based green system for exercise activity recognition and quality evaluation. Different from traditional features exploited in previous activity recognition works, like mean or max value, containing either time or frequency information, we select CSI-waveform shape as activity feature, which contains both time and frequency information. Human activities generally comprise of low-frequency components, while noise consists of high-frequency components. Instead of using common filtering methods to eliminate high-frequency noise, namely low-pass filter, we design a series of de-noise methods, including low-pass filtering, Principal Component Analysis (PCA) and median filtering. PCA can not only remove the in-band noise that traditional low-pass filter cannot eliminate, but also preserve activity-related information. Median filtering can further smooth the waveforms, facilitating activity recognition. Existing activity recognition systems usually work on one single motion whereas most exercise activities are repeated ones. Thus the First-order Difference (FOD) and the Fast Fourier Transform (FFT) are combined to detect activity patterns, and then each segment contains a complete CSI-waveform shape. SEARE primarily consists of three modules: 1) noise removal: applying a series of de-noise methods to original signals, including low-pass filtering, PCA, and median filtering; 2) activity feature extraction: segmenting CSI data and adopting the shape of CSI waveform as activity feature; 3) activity recognition: matching the unknown activity feature with pre-constructed standard activities features and selecting the best-fit one as recognized activity. Simultaneously we can make quality evaluation for these motions according to outputs  $D$  with a predefined threshold  $TH$ . Specifically, the motion is good when  $D < TH$  or bad when  $D > TH$ . The value of  $TH$  is obtained from a number of priori experiments, including standard and bad exercise motions. It is worth noting that in terms of green energy management for systems, SEARE will enter the sleep mode if there are no human activities in the monitor area for two minutes. The chief goals of our system are as follows: 1) being easily deployed, low-cost, and energy-efficient; 2) being able to accurately recognize some common exercise activities; 3) being feasible in both LOS and NLOS scenarios. Daily exercise activities in general are characterized by repeated motions. Thus, in this work, we only focus on four common and representative activities, namely dumbbell lifting, deep squat, kicking and boxing.

In view of the design goals mentioned above, there are five challenges.

The first challenge is that raw signals contain much noise and traditional filters such as low-pass filter fail to eliminate in-band noise, though they are able to remove pulse and burst noise. This paper adopts PCA to remove the in-band noise. It also preserves activity-related information and reduces computation complex.

The second challenge is that we discard all the other components but the first principal component in PCA, which captures the human movement signals along with noise. So that simple use of the first principal component may lead to fallacy. In our work, a series of de-noise methods are devised before and after PCA, including low-pass and median filtering, greatly reducing the noise in the first principal component.

The third challenge is how to extract distinguishing features for different exercise activities. Based on the fact that different activities are different in both time and frequency domains, we choose CSI-waveform shapes as activities features, which contain both time and frequency information.

The fourth challenge is how to automatically detect the starts and ends of activities. We find that signals change randomly in time series in static environment because of noise. Once a person moves, the signals become related with upward or downward changes. And exercises are a group of periodic activities. So that the FOD and FFT are combined to detect starts and ends of activities.

The fifth challenge is how to calculate the distance between two feature vectors with different lengths in time series, because different activities seem to vary in time lengths. In this work, the Dynamic Time Warping (DTW) is selected to solve this problem.

To sum up, the main contributions of this paper are as follows:

- SEARE is a novel and energy-efficient solution using ubiquitous WiFi for exercise activities recognition, which provides users with health management during they exercise.
- Instead of using common filtering methods to eliminate noise of high frequency in previous works, namely low-pass filter, we design a series of denoise methods, including low-pass filtering, PCA and median filtering. PCA can not only remove the in-band noise that traditional low-pass filters cannot eliminate, but also preserve activity-related information. Median filtering can further smooth the waveforms.
- Different from the single motion in existing activity recognition systems, most exercise activities are repeated ones. A novel approach is proposed to detect the starts and ends of activities by combining the FOD and FFT, after which each segment contains a complete CSI-waveform shape.
- SEARE is a lightweight and energy-efficient system implemented on existing commercial WiFi devices, with no need for any extra specialized hardware equipment and nodes. More importantly, it is characterized by sleep mode when no activities are performed in the monitor area for two minutes.
- SEARE is feasible in both LOS and NLOS scenarios, with average recognition accuracies of 97.8% and 91.2% respectively.

The rest of this paper proceeds as follows: we first review the related works in Section 2 and some preliminaries are introduced in Section 3. Section 4 presents the detailed design of the system, followed by experimental settings and performance evaluation in Section 5. Finally we conclude our work and look forward to the future works in Section 6.

## 2 RELATED WORK

In general, the approaches to activity recognition can be generally divided into two categories: device-based and device-free. Device-based methods require users to wear dedicated hardware and they are constrained by external environments, such as the requirement of enough light and line-of-sight condition. As a result, they are inconvenient in some scenes, like swimming and bathing, and the dedicated hardware increases the cost. In this paper, wireless signals are adopted to recognize activities, which is ubiquitous, non-invasive and low-cost.

A variety of wireless signals are used in the device-free activity recognition, which mainly fall into four types as follows:

**Specialized Devices Based.** Fine-grained signals can be collected by specially designed hardware devices or software-defined radio (SDR), a device providing complex signal characteristics. WiSee uses USRP to measure the Doppler shift caused by human activities and is able to recognize nine different gestures [13]. AllSee designs a dedicated analog circuit to extract amplitude characteristics of received signals, enabling the recognition of a set of eight gestures within a distance of 2.5 feet [23]. WiHear makes use of directional antennas to obtain CSI changes due to lips movements to identify speech content, nevertheless it lacks of effective denoising methods [14]. Unlike methods mentioned above, the prototype of SEARE is implemented on existing commercial WiFi devices without extra hardware devices and has a series of de-noising methods, which is privacy-preserving to some extend and is energy-efficient.

**Radar Based.** Radar-based activities detection requires specialized radar signals, such as Frequency Modulated Continuous Wave (FMCW) and ultra-wideband signals. The radar can detect movement speed of each part of the body by Doppler information. WiTrack utilizes FMCW to track the movement of people behind the wall, with an accuracy of about 20cm [24]. In this paper, there is no need for specially designed signals to identify activities for we adopt ubiquitous WiFi signal with narrower bandwidth instead.

**RSSI Based.** RSSI, that is Receive Signal Strength Indicator, describes channel quality and is the superposition of multipath signals. It can be easily obtained by wireless communication technologies [25], [26]. [27]and [28] exploit RSSI signal to recognize activities, such as lying down, standing up, and walking, with average accuracy of 80%. WiGest is able to identify different in-air hand gestures around users' mobile devices utilizing the changes of RSSI signals and its average accuracy is 96%, which sheds light on the hands-free gesture-based interaction with mobile devices [29]. However, the localization accuracy of RSSI is limited for it is a coarse granularity description of signals and is time-unstable. Thus RSSI is not appropriate for fine-grained activity recognition. In this paper, fine-grained CSI instead of RSSI is selected to recognize activities, with average recognition accuracy of around 94.5%.

**CSI Based.** Recently CSI-based activity recognition attracts much attention and stimulates a great deal of applications. WiFall implements fall detection for a single person [30]; Electronic Frog Eye counts the crowd by regarding the WiFi signal reflected by people as a virtual antenna [31]; E-

eyes identifies home activities, like dishwashing and bathing, through CSI values [32]. Wu et al. proposes WiDir that measures the direction of human walking by Fresnel model, with median error of less than 10 degrees [33]. In [34], the author combines CSI-speed model and CSI-activity model to recognize a given activity. WiVi, mainly dealing with coarse gestures and activities, uses virtual multi-antenna to identify signal fluctuations caused by human motions [16]. WiFinger [3], [35] is able to identify micro gestures, such as nine digital gestures, zoom out, zoom in, etc., with an accuracy of more than 90%; Wang et al. investigates the effect of human's position and body orientation on respiration detection based on the finding that human's breathing will have an impact on the received signals [19]. Although previous works have made great breakthrough in various applications, exercise activity recognition, a noteworthy area, is ignored. Our work exploits ubiquitous CSI for activity recognition and quality evaluation and its average recognition accuracy is above 90% under low packet transmission rate.

## 3 PRELIMINARIES

### 3.1 Channel State Information (CSI)

In typical indoor environments, signals propagate through multiple paths, and each of them leads to different amplitude attenuation, phase shift, and time delay. OFDM technology is an innovative and bandwidth-efficient multi-subcarrier modulation scheme. It can effectively combat multipath propagation, enabling signals to be reliably received. Thus OFDM technology is widely used in wireless communication system. In OFDM, signals are transmitted over orthogonal frequencies, which are called subcarriers. Based on OFDM, CSI describes the channel properties of a communication link and is able to distinguish multipath propagation at the subcarriers level. Utilizing off-the-shelf commercial WiFi devices,  $N$  groups of subcarriers in the form of CSI can be collected for every packet, and says the CSI as:

$$H = [H(f_1), H(f_2), \dots, H(f_N)] \quad (1)$$

In this paper,  $N = 30$ . Each CSI contains both amplitude and phase information of a subcarrier:

$$H(f_k) = ||H(f_k)||e^{j\angle H(f_k)}, k \in [1, 30] \quad (2)$$

Where  $H(f_k)$  is the CSI with central frequency of  $f_k$ ,  $||H(f_k)||$  and  $\angle H(f_k)$  denote its amplitude and phase respectively. In terms of a lot of random noise in raw phase information and unsynchronized time clock between transmitter and receiver, we limit our study to the amplitude information of CSI. To recognize the activities in the monitor area, we collect CSI data continuously and  $K$  measurements within a specific time window is selected as the input for our system. The collected CSI set can be denoted as:

$$\mathbb{H} = [H_1, H_2, \dots, H_K] \quad (3)$$

Currently, CSI can be mainly extracted from two tools: Atheros CSI tools and Intel CSI tools. In this paper, Intel CSI tool is adopted for our system implementation.

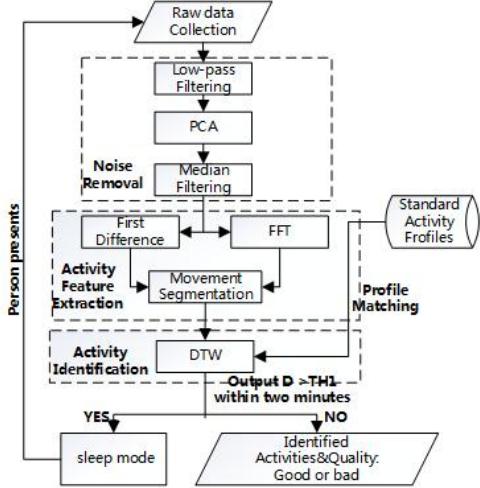


Fig. 1. SEARE system design overview.

### 3.2 Multiple Input and Multiple Output (MIMO)

MIMO technology, the core technology of the next generation of mobile communication, uses multiple antennas at the transmitter and receiver. The signals are sent and received through multiple antennas between the transmitter and receiver, which increases the communication capacity and improves the communication quality, without increasing the spectrum resources and the transmission power. Compared to a single antenna system which can only transmits or receives one data stream at a time, the MIMO system makes it possible that multiple data streams are transmitted and received simultaneously. So the MIMO technology is widely employed in wireless sensors network systems. Assuming that there are  $m$  and  $n$  antennas at the transmitter and receiver respectively, then all CSI data streams can be expressed as:

$$\begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{12} & \cdots & \tilde{H}_{1n} \\ \tilde{H}_{21} & \tilde{H}_{22} & \cdots & \tilde{H}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{m1} & \tilde{H}_{m2} & \cdots & \tilde{H}_{mn} \end{bmatrix} \quad (4)$$

Where  $\tilde{H}_{mn}$  represents the mn-th CSI data stream between the m-th transmit antenna and the n-th receive antenna, so the system has  $m \times n$  CSI data streams in total, each of which contains  $N$  ( $N = 30$ ) subcarriers and can be written as:

$$CSI = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1j} \\ H_{21} & H_{22} & \cdots & H_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ H_{i1} & H_{i2} & \cdots & H_{ij} \end{bmatrix}, i \in [1, mn], j \in [1, 30] \quad (5)$$

Where  $H_{ij}$  is the CSI value of the  $j$ -th subcarriers on the  $i$ -th CSI data stream. In this paper,  $m = 3, n = 3$ .

## 4 SYSTEM DESIGN

#### 4.1 Overview of System

SEARE principally consists of three modules: noise removal, feature extraction, and activity recognition, as shown in Fig. 1.

SEARE collects real-time CSI signals and passes them through Butterworth filter to remove high-frequency noise;

followed by PCA which can eliminate the in-band noise that Butterworth filter cannot do, and then the output signals are further smoothed by median filter.

To extract activities features, the FOD and FFT are combined to detect starts and ends of activities, so that each segment contains a complete CSI waveform. SEARE extracts CSI-waveform shapes of exercise activities as features.

To recognize activities, features database of standard activities should be built in advance, and then the similarities between the unknown CSI pattern and pre-constructed CSI-waveform shapes are calculated by DTW, outputting  $D$ . The best-fit one is regarded as recognized activity. Meanwhile according to the value of  $D$ , we can evaluate the activity quality with a predefined threshold  $TH$  and provide feedbacks for users.

To save energy, when there are no activities for two minutes, the system will enter the sleep mode, significantly cutting down the consumption of energy.

The main technical methods will be described in detail in the following section.

#### 4.2 De-noising Methods to CSI Data

### 4.2.1 Low-pass Filtering

Raw signals cannot be directly used for activity recognition for there are much noise, such as personnel interference, electromagnetic noise, etc. The main sources of noise in the CSI stream are from internal state transitions in WiFi NICs of transmitter and receiver, like transmission strength changes and transmission rate changes, which may produce pulses and burst noise. Fig. 2(a) and Fig. 2(b) are raw and filtered signals after Butterworth filter respectively when a person does exercise, where curves in different colors are the amplitude of different subcarriers. Specifically, there are 30 curves in each figure, because Intel 5300 device driver used in SEARE provides 30 groups of subcarriers. As seen, the original signals contain amounts of high-frequency noise, whereas human activities are in general low-frequency. Therefore, we adopt Butterworth filter to remove high-frequency noise and we empirically set the passband cutoff frequency to 30 Hz and the stopband cutoff frequency 80 Hz. On account of low sampling frequency, there are still plenty of in-band noise in filtered signals shown in Fig. 2(b). As a consequence, the filtered signals need further processing.

#### 4.2.2 PCA-based Filtering

It can be seen from Fig. 2(b) that CSI data changes due to human movements are relevant on all subcarriers, based on which we adopt PCA to eliminate the in-band noise and preserve activity-related information, as depicted in Fig. 2(c). As can be seen, the first principal component is superior to the other components in the description of activity characteristics, so SEARE discards all the other components but the first principal component. In a word, there are two advantages of PCA. First, PCA can reduce the computation complexity because it reduces the dimensionality of signals collected from 30 subcarriers in each TX-RX pair. Consequently, it automatically preserves the useful information related to activities. Second, PCA can remove the in-band noise that traditional low-pass filters fail to do, taking advantage that CSI variations caused by human movements are correlated

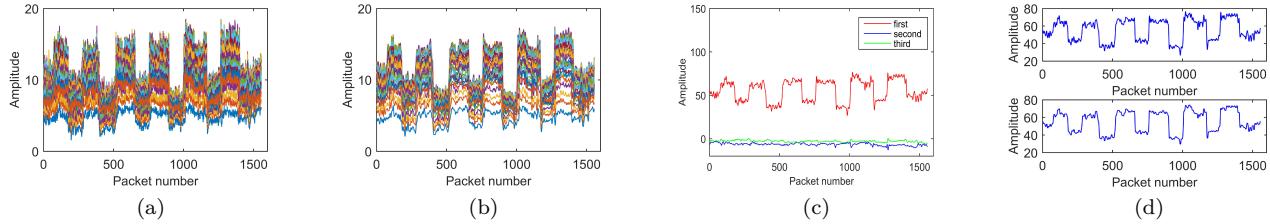


Fig. 2. The procedure of de-noising on raw CSI data. (a) Original CSI amplitudes of a exercise activity. (b) CSI amplitudes after Butterworth filtering. (c) CSI amplitudes after PCA. (d) CSI amplitudes after median filtering.

over all subcarriers in time series, while noisy components are not.

#### 4.2.3 Median Filtering

Fig. 2(d) plots the change caused by median filtering, consisting of two subfigures. The first one depicts the first principal component signals resulting from PCA. As can be seen, there are still some noisy components, which are likely to cause fallacious recognition. Hence we apply a median filter to further smooth the signals, facilitating the recognition of exercise activities. Finally, clean signals are obtained, as depicted in the second one.

### 4.3 Features Extraction

#### 4.3.1 CSI Data Segmentation

Considering that signals change randomly in time series in static environment where there are a large amount of random noise, while they become related and change upwards or downwards with the movement of person, and daily exercises are a group of periodic activities, we combine the FOD and FFT to detect the starts and ends of activities. The system requires a short interval between movements as a sentinel signal that can be identified. In detail, in order to weaken noisy interference, median filter is applied to the obtained clean signals. The input sequences of one-dimensional discrete signals and sliding window length of the median filter are denoted as  $\{x(1), x(2), \dots, x(n)\}$  and  $m$  respectively, where  $m$  is odd. Each time taking consecutive  $m$  data from the input signals sequences, the output signals can be written as  $\{x(i-v), \dots, x(i), \dots, x(i+v)\}$ , where  $v = (m-1)/2$ . Then the outputs are arranged in ascending order and the  $i$ -th one is selected as the final output. The median filter is expressed as follows:

$$y(i) = \text{Med}\{x(i-v), \dots, x(i), \dots, x(i+v)\} \quad (6)$$

Experiments show that the frequency response of the median filter is a irregular curve with a relatively flat mean. It is not largely fluctuant. Therefore we can draw that the value of transfer function for signals after median filter is approximately 1. In other words, the median filter has little effect on the frequency of signals and the spectrum is nearly unchanged.

In order to reduce the interference of external noise and facilitate the detection of the starts and ends of activities, we strengthen the frequency characteristics of signals by further processing output signals with the FOD in consideration of

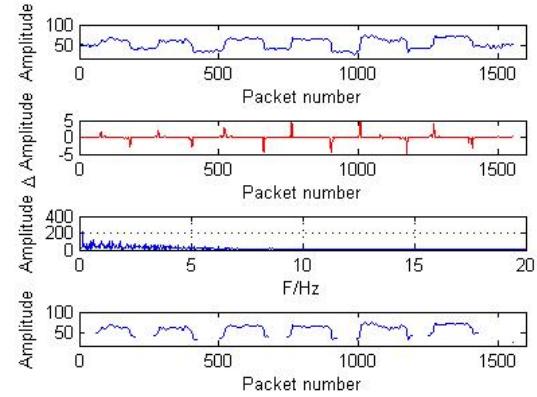


Fig. 3. The procedure of activity profile extraction.

the periodicity of exercises, which is the difference between two adjacent data as for one-dimensional signals  $y(t)$ :

$$Z(k) = \frac{\partial y}{\partial t} = y(k+1) - y(k), k = 0, 1, 2, \dots, M-1 \quad (7)$$

Finally, substituting the signal  $z(k)$  to FFT, written as

$$Z(u) = DFT(z) = \frac{1}{m} \sum_{k=0}^{M-1} z(k) e^{-j2\pi u k/M}, \quad u = 0, 1, 2, \dots, M-1 \quad (8)$$

From Equation (8) we can get the spectrum of the activity signals. The main spectrum  $f$  is selected as the activity frequency, and then the start and end points of the activity can be estimated. We denote these points as  $\{t_1^s, t_1^e, t_2^s, t_2^e, \dots, t_n^s, t_n^e\}$ , where  $[t_n^s, t_n^e]$  contains a CSI waveform of one activity. In the process of feature extraction, SEARE sets a guard interval  $T_b$  on both sides of  $t_n^s$  and  $t_n^e$  to ensure that  $[t_n^s, t_n^e]$  contains a complete CSI-waveform, where  $T_b = \frac{0.01}{f}$  in this paper. Consequently, the signal set becomes  $\{t_1^s - T_b, t_1^e + T_b, t_2^s - T_b, t_2^e + T_b, \dots, t_n^s - T_b, t_n^e + T_b\}$ . Fig. 3 shows the procedure of activity profiles extraction, consisting of four subfigures. The first one plots the clean CSI amplitude signals of dumbbell lifting after noise removal and the second one shows the output signals after FOD, where the peaks and the valleys represent the stars and the ends of actions respectively. The spectrum obtained by the FFT is illustrated in the third one, where the abscissa of the peak is the main frequency, which is used as the activity frequency. The last one is CSI-waveform shapes of the activity, where each CSI-waveform shape corresponds to a complete activity.

TABLE 1  
Feature values extracted from different exercise activities.

	Dumbbell lifting	Deep Squat	Kicking	Boxing
MEAN	64.87	55.71	26.14	54.77
MAX	74.63	65.65	31.59	57.85
MIN	10.67	14.31	16.06	43.87
MEDIAN	64.68	62.12	28.12	55.08
STD	8.80	15.10	4.41	2.28

#### 4.3.2 Feature Extraction

Traditional features such as the mean value (MEAN), max value (MAX), minimum value (MIN), and standard deviation (STD) cannot be used in this paper, because activities like dumbbell lifting and deep squat give almost the same values for these features, as shown in Table 1, which may fail in the classification of different exercise activities.

Based on the observation that different activities are different in both time and frequency domains, we choose CSI-waveform shapes as activity features, which contains both time and frequency information. Fig. 4 depicts features corresponding to four representative exercise activities, dumbbell lifting, deep squatting, kicking and boxing. As seen, they are absolutely different from each other, from which we can draw a conclusion that CSI-waveform shapes can be used as features to distinguish different exercise activities.

#### 4.4 Activity Recognition and Quality Evaluation

##### 4.4.1 Dynamic Time Warping

DTW is an algorithm based on dynamic programming for calculating the minimum distance between two data series, i.e.,  $X = (x_1, x_2, \dots, x_n)$  of length  $n$  and  $Y = (y_1, y_2, \dots, y_n)$  of length  $m$ , where  $m$  may not be equal to  $n$ . It allows a non-linear mapping from one data series to another by minimizing the distance between the two data series [15]. Hence DTW can work even if they are shifted or distorted. In contrast to DTW, traditional k-nearest neighbor algorithm (KNN) adopts Euclidean distance as measurement metric between data series, which fails in the distance calculation of two vectors not in the same length.

##### 4.4.2 Activity Recognition and Quality Evaluation

Because of the influence of the external factors on the data, the extracted feature vectors are processed by subtracting the mean value to obtain dimensionless data. We build the standard features database for individual activities in advance and then calculate the distance between the unknown activity feature and pre-constructed standard features to make a decision for activity recognition.

Considering that extracted feature vectors may be various in length, DTW is used for distance calculation, outputting  $D$ . The value of  $D$  describes the likelihood of two vectors. Specially, the larger the  $D$  is, the less similar the two vectors are. SEARE feeds the extracted features of unknown exercise activity to DTW and selects the best-fit one as identified activity.

At the same time, SEARE can further evaluate activities quality according to the result  $D$  by setting a predefined threshold  $TH$ . The value of parameter  $TH$  is obtained from

a plethora of priori experiments, including standard and bad exercise actions. In detail, if  $D > TH$ , the quality of this motion is bad and SEARE will remind the user that his or her exercise actions are not standard enough, providing the users with health assurance during they exercise.

It is worth noting that experiments show the value of  $D$  is much larger when there is no activity performance, compared with that when a user is doing exercise. Thereby, with a predefined threshold  $TH1$ , we can identify whether there is an exercise activity in the monitor area or not. The value of parameter  $TH1$  is out from a large number of priori experiments, including exercising and no activity scenes. If the value of  $D$  is larger than  $TH1$  for two minutes, the system will enter the sleep mode to save energy.

## 5 EXPERIMENT & EVALUATION

### 5.1 Implementation

#### 5.1.1 Experimental Setup and Platform

We prototype SEARE with existing commodity WiFi devices and evaluate its performance in a typical indoor environment, a 10m\*5m office room. Fig. 5 shows the experiment platform and area. Experimental platform includes a mini power machine, TP-Link router, three external antennas, LCD display and a number of notebooks, where the router and mini power machine with three antennas are as transmitter and receiver respectively and the distance between them is 2 meters. The transmitter operates in IEEE 802.11n AP mode at 2.4GHz with 20MHz bandwidth, and the receiver is equipped with Ubuntu system, virtual CSI, Intel 5300 wireless card and CSI Tool. In our experiments, the packet transmission rate is set to 50 packets per second and CSI data from each packet are recorded.

#### 5.1.2 Experimental Methodology

In order to establish standard features database for individual activities in advance, we have 5 specially trained volunteers do activities, standing in the middle position between the transmitter and the receiver. As mentioned above, in this work, we only focus on four representative activities, namely dumbbell lifting, deep squat, kicking and boxing, and each activity repeats 20 times. The system requires a certain time interval between successive movements. According to the methods introduced above, we can get 100 features for each activity, and then the standard features for these activities can be obtained by averaging them.

The test data can be divided into two categories: 1) data in line of sight (LOS) scenario: let a volunteer do exercise activities randomly standing in the middle position between the transmitter and the receiver on the LOS path, as depicted in Fig. 6(a); 2) data in non-line of sight (NLOS) scenario: let a volunteer randomly do exercise activities off the LOS path, as presented in Fig. 6(b). In order to facilitate research work, we collect the same number of data packets for each type of these activities.

#### 5.1.3 Evaluation Metrics

We mainly use the following metrics to evaluate the performance of SEARE.

- Confusion Matrix: each column shows the standard activity feature that is pre-established by our system

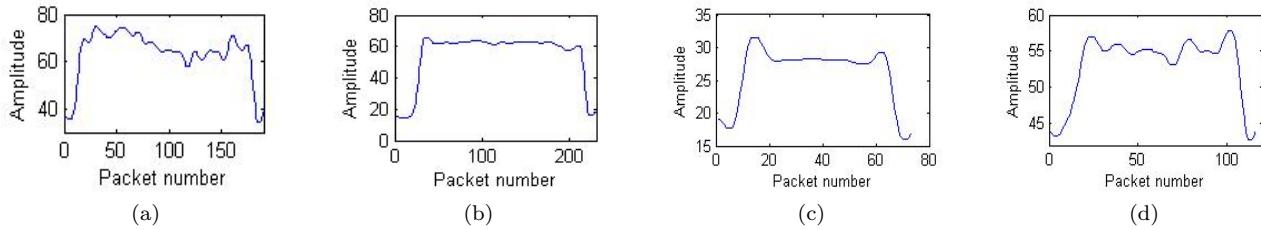


Fig. 4. Shape features of different exercise activities. (a) Shape of dumbbell lifting waveform. (b) Shape of deep squat waveform. (c) Shape of kicking waveform. (d) Shape of boxing waveform.

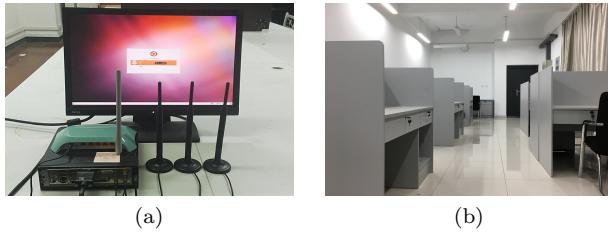


Fig. 5. Experiment platform and environment. (a) Experiment platform. (b) Experiment environment.

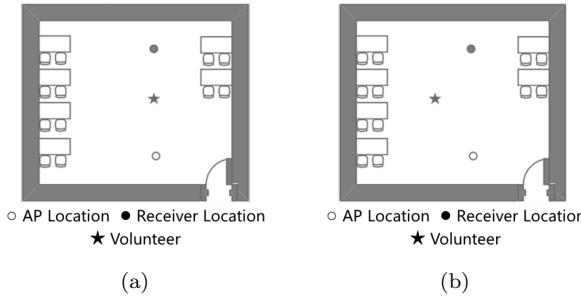


Fig. 6. Experimental areas. (a) LOS area. (b) NLOS area.

and each row represents under-test activity. Each cell in the confusion matrix is the distance  $D$  obtained by matching the corresponding activity feature in the row with the standard activity feature in the column by DTW.

- Recognition Accuracy: the percentage of the unknown exercise activities correctly recognized by our system.

## 5.2 Performance Evaluation

### 5.2.1 Overall Performance in LOS Scenario

In this paper, we carry out verification experiments. The extracted features for four under-test activities, dumbbell lifting, deep squat, kicking, and boxing, are denoted as Test1, Test2, Test3, and Test4 respectively. Matching them with the standard activities features and the results are illustrated in Table 2. As can be seen, compared to the distance between Test1 and the other three standard features, that between Test1 and the standard dumbbell lifting feature is the smallest, so that Test1 is recognized as dumbbell lifting; the distance between Test2 and the standard deep squat feature is the smallest compared to that between Test2 and the other three standard features, so that Test2 is recognized as deep squat; Test3 is kicking and Test4 is boxing by

TABLE 2  
Confusion matrix of distance between unknown and standard activities.

	Dumbbell lift	836.7*	9637.4	2040.0	2341.2
Dumbbell lift	836.7*	9637.4	2040.0	2341.2	
Deep Squat	17541.0	161.9*	2308.8	1138.4	
Kicking	184840.0	4465.6	38.5*	4041.4	
Boxing	1310.0	13755.0	1333.6	593.3*	
Test 1	Test 2	Test 3	Test 4		

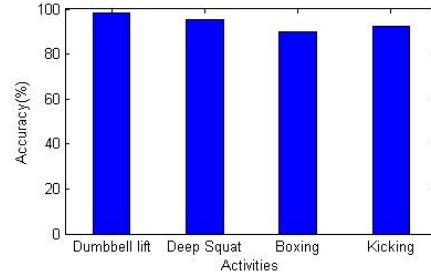


Fig. 7. Recognition accuracies of different exercise activities.

analogy. All the under-test activities are accurately identified. A large number of experiments demonstrate that SEARE can effectively recognize different exercise activities in LOS scenario, with an average accuracy of 97.8%.

Fig. 7 shows the recognition accuracies of different activities, from which we can see that the accuracies of deep squat and dumbbell lifting are much higher than those of boxing and kicking, because there is a Fresnel zone along TX-RX link, a series of concentric elliptical regions with alternating reinforced and weakened strength of a wave's propagation [19], [33]. Human activities passing through the Fresnel zone will lead to the amplitude shift. In terms of the shape of the Fresnel zone, signal variations caused by activities in the same vertical plane with TX-RX link, like deep squat and dumbbell lifting, are larger than activities in the other plane, like kicking and boxing. So the accuracies of deep squat and dumbbell lifting are higher than the other two activities.

### 5.2.2 Impact of NLOS

We are interested in the performance of SEARE off the LOS path, so let volunteers do exercise activities randomly in NLOS scenario. Collected data are processed by the same methods described in the former sections. The average recognition accuracy is around 91.2%, which validates good performance of SEARE in NLOS scenario and indicates

TABLE 3  
Distance between nonstandard and standard activities.

	Dumbbell lift	1.7E+04*	5.0E+03*	2.3E+03*	3.4E+03	2.6E+03	3.2E+03
	Deep Squat	6.6E+04	3.5E+04	1.5E+04	2.0E+03*	1.7E+03*	1.1E+03*
	Kicking	2.8E+05	2.5E+05	2.2E+05	3.3E+03	3.0E+03	3.1E+03
	Boxing	1.0E+05	7.2E+04	7.1E+04	6.5E+03	4.8E+03	6.1E+03
	Dumbbell lift 1	Dumbbell lift 2	Dumbbell lift 3	Deep Squat 1	Deep Squat 2	Deep Squat 3	

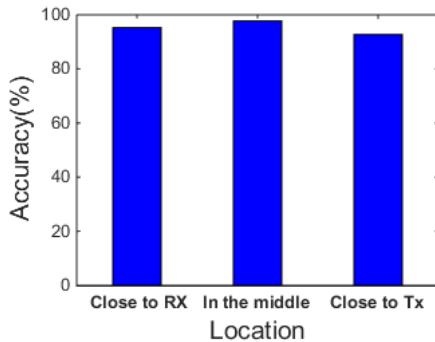


Fig. 8. Recognition accuracies of different locations.

the feasibility of SEARE in various scenarios, not requiring human to stand directly on the LOS path.

### 5.2.3 Impact of Locations

Given volunteers' facial orientations are parallel to the TX-RX link and experiments are conducted in the LOS scenario. Let them do exercise activities at different locations, namely close to the receiver, in the middle position between the transmitter and the receiver, and close to the transmitter. The collected data are processed as the methods mentioned before and recognition accuracies are drawn in Fig. 8. As seen, SEARE performed well no matter where the user stands, providing recognition accuracies of above 90%. When the user is close to the transmitter or receiver, the recognition accuracy becomes slightly lower, because he or she hinders signals propagation.

### 5.2.4 Impact of Body Directions

Given the experiments are conducted in the LOS scenario and the volunteers are standing in the middle position between the transmitter and the receiver, let them change their body orientation, namely 30° towards the TX-RX link, 90° towards the TX-RX link, and 180° towards the TX-RX link. The collected data are processed by the methods mentioned before and the recognition accuracies are shown in Fig. 9. As seen, the accuracy when users stand with 90° towards the TX-RX link is higher than that with other angles, because the number of elliptical regions in the Fresnel zones through which exercise activities pass with 90° towards the TX-RX link is much more than that with other angles do. As a result, signals changes resulted from exercise activities with 90° towards the TX-RX link are much more significant than those with the other angles. The accuracies of activity recognition when users stand 30° and 150° towards the TX-RX link are approximate, because in term of the ellipsoid Fresnel zone, 30° and 150° towards TX-RX link are symmetrical, and then the effect

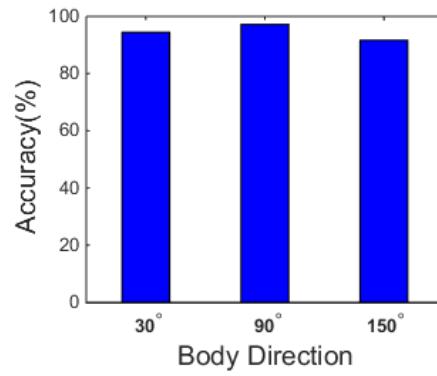


Fig. 9. Recognition accuracies of different body direction.

on received signals caused by those activities are almost the same.

### 5.3 Quality Evaluation

Table 3 shows part of experimental results of non-standard activities, where each column shows the standard activity feature and each row represents an under-test activity. Compared to the results in Table 2, the distance  $D$  in Table 3 is much larger, which indicates that the activities are not standard to a certain degree and they are likely to be harmful to users' health. Therefore according to the results, we can make quality evaluation by setting a predefined threshold  $TH$ . If  $D > TH$ , SEARE will remind the user to normalize activities.

## 6 CONCLUSION & FUTURE WORK

In this paper, we present the design and implementation of SEARE, a novel energy-efficient solution using ubiquitous WiFi for exercise activities recognition. In order to improve the reliability of the system, we design a series of de-noising methods, including low-pass filtering, PCA, and median filtering. A novel method is proposed to detect the starts and ends of each action by combining the FOD and FFT. Then CSI segments are obtained and each segment contains a complete CSI-waveform shape, which is used as activity feature. Finally, DTW is adopted to calculate the distance between unknown features and standard features and meanwhile SEARE can make a quality evaluation according to outputs  $D$  with a predefined threshold  $TH$ . The prototype of SEARE is implemented in the typical indoor environment and experimental results show the feasibility of SEARE in various scenarios. Our system is a energy-saving scheme, requiring no specialized hardware facilities and

characterized by sleep mode, shedding light on practical in-home exercise monitoring. Currently we limit our system to activity recognition in case of individual person presence. In the future, we will further study the scene of multiple targets. When there are a number of people in the monitoring area, the received signal is the superposition of multiple signals reflected by multiple individuals. We may then need to devise an approach to distinguish different CSI waveforms caused by different individuals, which is indeed a big challenge.

## References

- [1] Z. Zhou, C. Wu, Z. Yang and Y. Liu, "Sensorless sensing with WiFi," in Tsinghua Science and Technology, vol. 20, no. 1, pp. 1-6, Feb. 2015.
- [2] C. Zhang, F. Li, J. Luo, et al., "iLocScan: Harnessing Multipath for Simultaneous Indoor Source Localization and Space Scanning," in Proc ACM SenSys, 2014, pp. 91-104.
- [3] H. L, W. Yang, J. Wang, et al., "WiFinger: Talk to your smart devices with finger-grained gesture," in Proc. ACM UbiComp, 2016, pp. 250-261.
- [4] S. Gupta, D. Morris, S. Patel, and D. Tan, "Soundwave: Using the doppler effect to sense gestures," in Proc. ACM CHI, 2012, pp. 1911-1914.
- [5] M. Scholz, T. Riedel, M. Hock, and M. Beigl, "Device-free and device-bound activity recognition using radio signal strength," in Proc. ACM AH, 2013, pp. 100-107.
- [6] P. W. Juan, M. Kolsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications," CACM., vol. 15, no. 10, pp. 60-71, Feb. 2011.
- [7] Microsoft Kinect. <http://www.roborrealm.com/help/Microsoft-Kinect.php>
- [8] S. Nirjon, J. Gummesson, D. Gelb, et al., "Typingring: A wearable ring platform for text input," in Proc. ACM MobiSys, 2015, pp. 227-239.
- [9] A. Parate, M.-C. Chiu, et al., "Risq: Recognizing smoking gestures with inertial sensors on a wristband," in Proc. ACM MobiSys, 2014, pp. 149-161.
- [10] J. Wang, D. Vasishth, and D. Katabi, "Rf-idraw: virtual touch screen in the air using rf signals," in Proc. ACM SIGCOMM, 2014, pp. 235-246.
- [11] F. Xiao, Z. Wang, N. Ye, R. Wang, X. Li, "One More Tag Enables Fine-Grained RFID Localization and Tracking," IEEE/ACM Transactions on Networking, DOI:10.1109/TNET.2017.2766526. 2017:1-14.
- [12] Z. Wang, F. Xiao, N. Ye, R. Wang, P. Yang, "A See-through-wall System for Device-free Human Motion Sensing Based on Battery-free RFID," ACM Transactions on Embedded Computing Systems, vol. 17, no. 1, pp.1-21, 2017.
- [13] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in Proc. ACM MobiCom, 2013, pp. 27-38.
- [14] G. Wang, Y. Zou, Z. Zhou, et al., "We can hear you with wi-fi!" in Proc. ACM MobiCom, 2014, pp. 593-604.
- [15] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using wifi signals," in Proc. ACM MobiCom, 2015, pp. 90-102.
- [16] F. Adib, and D. Katabi, "See through walls with wifi," in Proc. ACM SIGCOMM, 2013, pp. 75-86.
- [17] W. He, K. Wu, Y. Zou and Z. Ming, "WiG: WiFi-Based Gesture Recognition System," in Proc. ACM ICCCN, Las Vegas, NoV, 2015, pp. 1-7.
- [18] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu and J. Cao, "Non-Invasive Detection of Moving and Stationary Human With WiFi," IEEE Trans. Selected Areas in Communications., vol. 33, no. 11, pp. 2329-2342, Nov. 2015.
- [19] H. Wang, D. Zhang, J. Ma, et al., "Human respiration detection with commodity wifi devices: Do user location and body orientation matter?" in Proc. ACM UbiComp, 2016, pp. 25-36.
- [20] J. Liu, Y. Wang, Y. Chen, et al., "Tracking Vital Signs During Sleep Leveraging Off-the-shelf WiFi," in Proc. ACM MobiHoc, 2015, pp.267-276.
- [21] X. Liu, J. Cao, S. Tang, J. Wen and P. Guo, "Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices," IEEE Trans. Mobile Comput., vol. 15, no. 10, pp. 2466-2479, Oct. 2016.
- [22] H. Zhu, F. Xiao, L. Sun, R. Wang and P. Yang, "R-TTWD: Robust Device-Free Through-The-Wall Detection of Moving Human With WiFi," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 5, pp. 1090-1103, May. 2017.
- [23] B. Kellogg, V. Talla, and S. Gollakota, "Bringing gesture recognition to all devices," in Proc. ACM NSDI, 2014, pp. 303-316.
- [24] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in Proc. ACM NSDI, 2014, pp.317-329.
- [25] F. Xiao, W. Liu, Z. Li, L. Chen, and R. Wang, "Noise-Tolerant Wireless Sensor Networks Localization Via Multi-norms Regularized Matrix Completion," IEEE Transactions on Vehicular Technology, DOI:10.1109/TVT.2017.2771805. 2017:1-10.
- [26] J. W. Zhang, C. Yin, C. Song, R. Liu and Bo Li, "Numerical Simulation and Experiments on Mono-polar Negative Corona Discharge Applied in Nanocomposites," IEEE Transactions on Dielectrics and Electrical Insulation, 2017, vol. 24, no. 2, pp. 791-797.
- [27] 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. MoMM, 2013, pp. 43-52.
- [28] S. Sigg, M. Scholz, S. Shi, Y. Ji and M. Beigl, "RF-Sensing of Activities from Non-Cooperative Subjects in Device-Free Recognition Systems Using Ambient and Local Signals," IEEE Trans. Mobile Comput., vol. 13, no. 4, pp. 907-920, Apr. 2014.
- [29] H. Abdelnasser, M. Youssef and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in Proc. IEEE INFOCOM, 2015, pp. 1472-1480.
- [30] C. Han, K. Wu, Y. Wang and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," in Proc. IEEE INFOCOM, 2014, pp. 271-279.
- [31] W. Xi et al., "Electronic frog eye: Counting crowd using WiFi," in Proc. IEEE INFOCOM, 2014, pp. 361-369.
- [32] Y. Wang, J. Liu, Y. Chen, et al., "E-eyes: Device-free location-oriented activity identification using fine-grained wifi signatures," in Proc. ACM MobiCom, 2014, pp.617-628.
- [33] D. Wu, D. Zhang, C. Xu, et al., "WiDir: Walking direction estimation using wireless signals," in Proc. ACM UbiComp, 2016, pp. 351-362.
- [34] W. Wang, A. X. Liu, M. Shahzad, et al., "Understanding and Modeling of WiFi Signal Based Human Activity Recognition," in Proc. ACM MobiCom, 2015, pp. 65-76.
- [35] S. Tan, J. Yang, "WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition," in Proc. ACM MobiHoc, 2016, pp. 201-210.



Fu Xiao (M'12) received the Ph.D. degree in computer science and technology from the Nanjing University of Science and Technology, Nanjing, China, in 2007. He is currently a Professor and a Ph.D. Supervisor with the School of Computers, Nanjing University of Posts and Telecommunications. His main research interests are wireless sensor networks and mobile computing. He is a member of the IEEE Computer Society and the Association for Computing Machinery.



Jing Chen received the B.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2012, where she is currently pursuing the M.E. degree in computer application technology. Her current research interests are wireless sensor networks. E-mail: m18020129701@163.com.



Xiaohui Xie received Bachelor in Communication Engineering from Nanjing University of Posts and Telecommunications, China. He is currently postgraduate in School of Computer, Nanjing University of Posts and Telecommunications, China. His main research interest is Delay Tolerant Networks. E-mail: xiexh\_njupt@163.com



Lijuan Sun received Ph.D in Information and Communication from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2007. She is currently a Professor and Ph.D supervisor with the School of Computer, Nanjing University of Posts and Telecommunications, Nanjing, China. Her main research interests are wireless sensor networks and wireless mesh networks.



Linqing Gui received Ph.D. degree in information science from INSA-Toulouse, France, in 2013. He is currently a teacher with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. His main research interests are wireless sensor networks, mobile computing and mobile communication.



Ruchuan Wang received the B.S. Degree in Calculation Mathematics from Information Engineering University, Beijing, China in 1968. He is currently a Professor and PhD supervisor with the School of Computer, Nanjing University of Posts and Telecommunications, Nanjing, China. His main research interests are Wireless Sensor Networks, information security and computer software.