WiStep: Device-free Step Counting with WiFi Signals

YANG XU, WEI YANG*, JIANXIN WANG, XING ZHOU, HONG LI, and LIUSHENG HUANG,
University of Science and Technology of China, China

Inspired by the emerging WiFi-based applications, in this paper, we leverage ubiquitous WiFi signals and propose a device-free step counting system, called WiStep. Based on the multipath propagation model, when a person is walking, her torso and limbs move at different speeds, which modulates wireless signals to the propagation paths with different lengths and thus introduces different frequency components into the received Channel State Information (CSI). To count walking steps, we first utilize time-frequency analysis techniques to segment and recognize the walking movement, and then dynamically select the sensitive subcarriers with largest amplitude variances from multiple CSI streams. Wavelet decomposition is applied to extract the detail coefficients corresponding to the frequencies induced by feet or legs, and compress the data so as to improve computing speed. Short-time energy of the coefficients is then calculated as the metric for step counting. Finally, we combine the results derived from the selected subcarriers to produce a reliable step count estimation. In contrast to counting steps based on the torso frequency analysis, WiStep can count the steps of in-place walking even when the person's torso speed is null. We implement WiStep on commodity WiFi devices in two different indoor scenarios, and various influence factors are taken into consideration when evaluating the performance of WiStep. The experimental results demonstrate that WiStep can realize overall step counting accuracies of 90.2% and 87.59% respectively in these two scenarios, and it is resilient to the change of scenarios.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; *Empirical studies in ubiquitous and mobile computing*;

Additional Key Words and Phrases: Channel State Information (CSI), Device-free Step Counting, WiFi Signals

ACM Reference Format:
Yang Xu, Wei Yang, Jianxin Wang, Xing Zhou, Hong Li, and Liusheng Huang. 2017. WiStep: Device-free Step Counting with WiFi Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 172 (December 2017), 23 pages. https://doi.org/10.1145/3161415

1 INTRODUCTION

In 2014, more than 1.9 billion adults aged 18 years and older were overweight. Of these over 600 million adults were obese. The worldwide prevalence of obesity more than doubled between 1980 and 2014[1]. Despite dietary adjustment, pharmacological and bariatric surgical approaches, daily exercise is an efficient treatment of obesity. Step counters are currently becoming popular among young people as an exercise

*This is the corresponding author.
[1]Obesity and overweight fact sheet, http://www.who.int/mediacentre/factsheets/fs311/en/

Authors' address: Yang Xu, smallant@mail.ustc.edu.cn; Wei Yang, qubit@ustc.edu.cn; Jianxin Wang, jacywang@mail.ustc.edu.cn; Xing Zhou, zhou999@mail.ustc.edu.cn; Hong Li, sa514012@mail.ustc.edu.cn; Liusheng Huang, lshuang@ustc.edu.cn, University of Science and Technology of China, School of Computer Science and Technology, Hefei, Anhui, 230027, China.

2474-9567/2017/12-ART172 $15.00
https://doi.org/10.1145/3161415

measurer and motivator, since step count can give people encouragement to compete with themselves in getting fit and losing weight [35]. Step count is also considered as an important parameter in the indoor localization [8, 17] and gait recognition [38, 43].

Traditional step counting methods continuously monitor the acceleration and angular variations of the accelerometer and gyroscope attached on a human body to calculate her walking step count [9, 11, 25]. However, the placement of the devices dominates the result of step counting, which is a major cause of overcounting or undercounting. In order to remedy the defect, some vision-based methods [7, 24] are proposed, which record step count by identifying the appearance and disappearance of a user's feet in the camera range of a mobile phone. Although these methods are resilient to some non-walking fluctuations, they are sensitive to the light condition and can't work in darkness, and the always-on cameras are power-consuming for smartphones and infringe on users' privacy. Moreover, the above methods require users to carry or wear essential devices (*e.g.*, mobile phones or wristbands), which may be uncomfortable for some special groups, such as children or elders, and may lose efficacy when users forget to wear the devices. Therefore, we are inspired to develop a device-free step counting system in this paper.

Owing to the Doppler shift and multipath effect of radio signals [34], different human activities can introduce different Doppler shifts and multipath distortions in the wireless signals without any help of wearable devices. Thus, many researchers have exploited these characteristics of wireless signals to develop various device-free human activity recognition systems [3, 19, 20, 26, 36, 39]. For example, WiSee is devised to recognize 9 gestures (*e.g.*, kicking, bowling) using the Doppler shifts, which are caused by body movements, in WiFi signals [26]. RT-Fall is proposed to detect the falling movement using phase differences of Channel State Information (CSI) of WiFi signals [37]. Meanwhile, WiHear can sense the variations of the CSI induced by lips and recognize the spoken words [36]. Wi-Sleep [21] and WiBreathe [30] both leverage wireless signals to monitor a user's respiration rate but with different deployments. In addition, our previous work WiFinger is designed to realize finger-grained gesture recognition and continuous text input [20]. Generally, most of the radio-based human activity recognition systems focus on the question "*Is this the target movement*?" Every movment is labeled with a classification result in terms of its feature vector, which means that these systems can only recognize "walking" rather than counting the steps.

Apart from recognizing activities, radio signals are also applied to extract human gait information. For example, WiWho [43] and WifiU [38] can extract the time-domain and frequency-domain gait features in WiFi signals to identify different people, and they both have the potential to count walking steps. However, WiWho and WifiU are assumed to work in a confined spaces like corridor or narrow entrance, and the target person needs to walk along a straight path in a predefined direction, which are inapplicable to general step counting scenarios. Moreover, WifiU mainly reserves the steady phase within a walking period to estimate gait cycle time, thus the steps in the acceleration and deceleration phases are omitted. WiGait [15] is implemented without the above constraints and can achieve high accuracy of gait velocity estimation in contrast to the VICON motion tracking system. However, it ignores all in-place activities which restrains it from recording the steps of activities like in-place walking/running. Besides, WiGait is built upon WiTrack2.0 [1] which uses an expansive software radio platform (thousands of dollars compared to tens of dollars for a commercial WiFi router) and the specialized Frequency-Modulated Continuous-Wave (FMCW) radar to detect human movements.

Considering the ubiquity of WiFi devices in our daily lives, it is meaningful to study and build a customized step counting system using the pervasive WiFi signals. Consequently, we develop WiStep, which is a WiFi-based device-free step counting system. Here, we care more about the question "*To what extent can WiFi be utilized to count walking steps*?" The key insight of WiStep is that when a person is walking, her torso and limbs move at different speeds, which modulates WiFi signals to the

propagation paths with different lengths and introduces different frequency components into the received CSI measurements. In WiStep, we first select the segments involving walking movements, and then dynamically pick out the subcarriers with largest variances from multiple CSI streams. The wavelet decomposition is conducted to compress and split the segments into multiple layers that correspond to different frequency ranges. The detail coefficients, which mainly represent the speed of legs, are reserved to construct a step profile. Subsequently, the short-time energy of the coefficients is calculated to analyze and get the step count derived from each subcarrier. Finally, we combine multiple results of the selected subcarriers to ensure a reliable output.

The main contributions of this work are summarized as follows:

- Based on the multipath propagation model, we focus on the frequency components, which are mainly induced by a person's feet or legs when she is walking, in CSI measurements, and propose WiStep, a device-free step counting system using WiFi signals. Compared with existing work which extracts gait features based on the torso frequency estimation, WiStep can count the steps of in-place walking even when the torso speed of the person is almost null.
- Two critical techniques are proposed to reduce the computing resources and guarantee a reliable result: (i) for detecting a walking movement, we utilize a threshold-based method with a simple but efficient metric termed Weighted Energy of Interest (WEI); (ii) we dynamically select the subcarriers with large variance from multiple CSI streams to construct the step profile and further calculate and calibrate the final step count.
- We implement WiStep on commodity WiFi devices and evaluate its performance with 50 volunteers in two different indoor environments. The experimental results demonstrate that WiStep can realize overall step counting accuracies of 90.2% and 87.59%, respectively, in these scenarios, and it is robust to the change of wireless environments.

2 RELATED WORK

Our work is related to previous arts of traditional step counting applications and some emerging radio-based applications.

2.1 Traditional Step Counting Applications

Traditional step counting methods typically employ sensors (*e.g.*, accelerometer, gyroscope) [8, 11, 17, 28] and cameras [7, 24] to count users' walking steps.

Sensor-based step counting systems generally leverage accelerometers and gyroscopes to collect users' movement information, such as accelerations. Modern smart devices, such as smartphones and smartwatches, are basically equipped with accelerometers and gyroscopes, which enable their ability of counting steps [6, 9, 25, 32, 42]. Kim and Kwon [18] tried to employ some choreographed thresholds to obtain the sensory data induced by walking and further estimated the step count with smartphones. Seo et al. [32] utilized an appropriate threshold to detect walking activity and introduced an advanced zero crossing technique for step counting. Oliver and Flores-Mangas [32] and Capela et al. [32] fixed the sensors to the user's chest and waist respectively to eliminate the fluctuations caused by hands or arms. Pan and Lin [32] designed a series of variables to optimize the movement segmentation and used correlation coefficients to identify every step.

Vision-based methods count steps by identifying the appearance and disappearance of a person's feet in the camera range of a mobile phone. For example, Aubeck et al. [7] matched the captured photos with the foot templates to detect steps, but the template matching often loses efficacy with fuzzy photos distorted by the user's fast moving. Ozcan and Velipasalar [24] exploited K-means clustering to the extracted

features of an image and traced the trajectories of the cluster center to recognize the user's steps reliably. However, the vision-based methods have many non-negligible drawbacks, such as light condition, power consuming and personal privacy. Moreover, users must hold the smartphones and let the cameras face their feet all the time, which is inconvenient and unpractical for users in practice. The above methods all require users to carry or wear essential devices, which may be uncomfortable for children or elders and may lose efficacy when users forget to wear the devices.

Different from these methods, our WiStep is designed to be device-free.

2.2 Emerging Radio-based Applications

The emerging radio-based applications are widely applied to human activity recognition [14, 19, 26, 40], indoor localization [1, 31, 41], gait recognition [15, 38, 43], entertainment [27] and so on. Here, we introduce some work in the areas of human activity recognition and gait recognition, which are more related to ours.

Human Activity Recognition: By extracting the micro-Doppler signatures of the wireless signals reflected off a human body, Kim and Ling [19] recognized 7 human activities using a continuous-wave Doppler radar. Adib et al. [1, 2] developed WiTrack and WiTrack2.0, which leverage the FMCW radar to track human movements behind the wall. Based on WiTrack, Adib et al. [4] devised Vital-Radio to simultaneously monitor the breathing and heart rates of multiple people. Compared with the broad bandwidth of radar signals (*e.g.*, in [2], the bandwidth of FMCW signals is 1.79 GHz), WiFi signals have much narrower bandwidth (usually 20 or 40 MHz), which results in low time resolution when separating different propagation paths [41].

As for WiFi-based systems, some researchers utilized the specialized devices, *i.e.*, Software Defined Radios (SDRs), to generate WiFi signals and analyze the echoes. For example, Adib and Katabi [3] employed the Universal Software Radio Peripheral (USRP) and developed Wi-Vi to sense the relative locations of multiple moving people through walls. Pu et al. [26] proposed WiSee that uses the USRPs to extract small Doppler shifts in WiFi signals and recognize 9 human gestures. Ravichandran et al. [30] proposed WiBreathe that utilizes two USRPs equipped with directional antennas as the transmitter and the receiver to estimate a person's respiration rate. Since SDRs can be programmed to control the transmission power and even modify the WiFi protocols [26], they guarantee more reliable communication links and more flexible operations than commercial WiFi devices, *e.g.*, wireless routers, but SDRs are too expensive to be widely deployed in practice.

Recently, with the tool released in [13], Channel State Information (CSI) can be obtained from commercial WiFi Network Interface Cards (NICs), *e.g.*, Intel 5300 NIC. Subsequently, Wang et al. [39] proposed CARM, a CSI-based human activity recognition and monitoring system, and introduced two specific models to quantify the relation between measured CSI values and human activities. Han et al. [14] and Wang et al. [37] separately used CSI amplitudes and phase differences to detect the falling movment. Wang et al. [36] proposed WiHear which exploits the CSI variations induced by lip movements to recognize people's words. Ali et al. [5] utilized CSI value changes and designed WiKey to identify different keystrokes. Liu et al. [21] explained that the CSI amplitudes could be used to sense the rhythmic patterns of respiration while the CSI phase didn't show a distinct correlation with breathing. Li et al. [20] developed WiFinger to recognize fine-grained finger gestures and realize continuous text input.

However, most of these systems are designed to map the extracted features to different target movments, *e.g.*, falling, walking and running, and can't be used to count walking steps.

Gait Recognition: With radar signals, Tahmoush and Silvious [33] extracted the micro-Doppler features, *e.g.*, torso speeds, stride rates, as the gait characteristics for different people, and achieved a

maximal gait recognition accuracy of around 80%. Hsu et al. [15] devised WiGait, which is built upon WiTrack2.0, to continuously monitor the gait velocity and stride length of a person. However, it ignores all movements whose ranges are less than 4 meters, which limits its ability to count the steps of activities like walking in place.

By utilizing the time-domain and frequency-domain gait features extracted from CSI, Zeng et al. [43] and Wang et al. [38] separately built systems called WiWho and WifiU to identify different people. WiWho mainly focuses on the low frequency band 0.3∼2 Hz of CSI, which contains much interference induced by slight body movements and environment changes. This hinders WiWho from working when the person is as far as more than 1 meter from the Line-Of-Sight (LOS) path of its sender and receiver. While WifiU illustrates that human walking activity can induce frequencies of 27∼60 Hz in CSI measurements. It conducts Principal Component Analysis (PCA) on 180 CSI subcarriers and keeps the first 20 PCA components to elaborate a synthetic spectrogram. Then the upper contour representing the torso reflections is outlined to estimate the gait cycle time in the period where only the steady phase of walking is reserved, which implies WifiU ignores the steps in the unsteady phases, *i.e.*, acceleration and deceleration phases.

In addition, although the gait cycle time (in WifiU) and the stride frequency (in WiGait) can be used as human-dependant features for gait recognition, they mainly reflect an average movement state of a person over a period of time. However, step count is more like an instantaneous metric, and all the steps should be recorded for an entire walking movement. To derive steps from the average metrics like stride frequency, the accurate walking time is needed. Considering that a person usually walks slower in the acceleration and deceleration phases than that in the stable phase, which makes the product of stride frequency and walking time larger than the actual step count.

Different from the above work, WiStep concentrates on the movements of feet or legs and is able to recognize the in-place walking. WiStep exploits the time- and frequency-domain characteristics of CSI and dynamically selects multiple sensitive subcarriers to construct a step profile, based on which a peak counting scheme is proposed to compute the steps within the entire walking period including the acceleration and deceleration phases.

3 PRELIMINARIES

To understand how walking and running movements exert impacts on wireless signals, we first give a brief overview of CSI. And then, we introduce the wireless propagation model which further facilitates the understanding of correlation between human activities and the CSI measurements.

3.1 Channel State Information

In wireless communication field, the frequency-domain Channel Frequency Response (CFR) characterizes the small-scale multipath effect and the frequency-selective fading of the wireless channels. Let $X(f,t)$ and $Y(f,t)$ separately denote the transmitted and received signal vectors, and the relation between X and Y can be modeled as

$$Y(f,t) = H(f,t) \times X(f,t) + \mathcal{N}_{noise}, \tag{1}$$

where f is the carrier frequency, t indicates that the channel is time-varying, $H(f,t)$ is the complex valued CFR and $\mathcal{N}_{noise}$ represents the noise.

In Orthogonal Frequency Division Multiplexing (OFDM) system, Channel State Information (CSI) is used to monitor the channel properties, and it is the simple version of CFR with discrete subcarrier frequency f_i $(i = 1, 2, ..., N_S$, where i is the subcarrier index and N_S denotes the total number of subcarriers across an OFDM channel). With the tool in [13], we can get a CSI measurement consisting

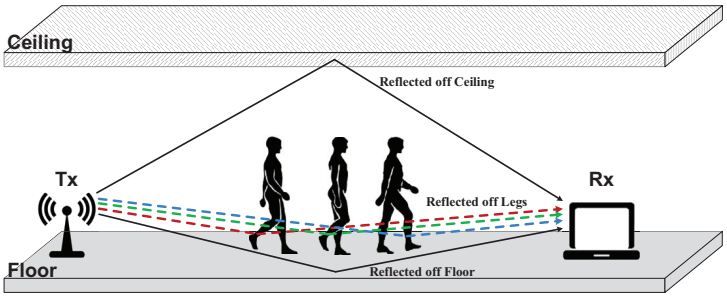


Fig. 1. Multipath propagation model in indoor environment.

of 30 matrices with dimensions $N_{Tx} \times N_{Rx}$ from one physical frame, where N_{Tx} and N_{Rx} represent the number of antennas of the transmitter (Tx) and the receiver (Rx), respectively. We regard the CSI measurements collected from each Tx-Rx antenna pair as a *CSI stream*, thus there are $N_C = N_{Tx} \times N_{Rx}$ time-series CSI streams.

3.2 Multipath Propagation Model

Fig. 1 simply illustrates the multipath propagation process of wireless signals traveling from a transmitter to a receiver in indoor environment. Without considering the reflections of walls and furniture, one transmitted signal can be reflected off the floor, the ceiling and the walking person and propagates through multiple paths before arriving at the receiver, which is called *multipath effect* [29]. The multiple paths can be divided into two categories: static paths (caused by the floor or other static reflectors as expressed by solid lines in Fig. 1) and dynamic paths (caused by human body as expressed by dotted lines in Fig. 1). If there totally exist N_P propagation paths among which N_D paths are dynamically varying, then $H(f_i, t)$ of the i^{th} subcarrier at time t can be expressed as

$$H(f_i, t) = \sum_{n=1}^{N_P} a_n(f_i, t) e^{-j2\pi f_i \tau_n(t)} = H_s(f_i) + \sum_{k=1}^{N_D} a_k(f_i, t) e^{-j2\pi f_i \tau_k(t)}, \tag{2}$$

where $H_s(f_i)$ is the sum of responses of the static paths[2], $a_k(f_i, t)$ represents the attenuation of the k^{th} dynamical path, and $2\pi f_i \tau_k(t)$ denotes the phase shift of the path k with a transmission delay $\tau_k(t)$ ($\tau_k(t) = d_k(t)/c$, where $d_k(t)$ is the path length and c is the light speed). Let λ_i be the wavelength of subcarrier i, and $\lambda_i = c/f_i$. Thus, $2\pi f_i \tau_k(t)$ can be rewritten as $2\pi d_k(t)/\lambda_i$, which means when the path changes λ_i, the phase of the received signal changes 2π accordingly. Besides, a λ_i displacement of the person can roughly cause a $2\lambda_i$ length change of the dynamical path k (round trip), which introduces 4π phase change. According to the principle of superposition of waves, the 4π phase change finally induces 2 cycles of the amplitude change of CSI values [39], which reveals an approximate relation between human moving speed V and frequency F of amplitude variation of CSI values, *i.e.*, $F = 2V/\lambda$, where λ is the wavelength of a certain subcarrier in OFDM system. Based on this relation, we find that if a person performs an in-place movement like walking on a treadmill, the speed of her torso is very low but her feet and legs still move at a relatively high speed, thus the torso-related gait features used in [15, 33, 38] won't work. In this paper, we decide to exploit the frequency components induced by feet or legs rather than torso to build our step counting system, and the detailed explanation will be given later.

[2] $H_s(f_i)$ can be regarded as a constant, since signals traveling through the static paths have relatively invariable path length, thus the propagation attenuation and transmission delay can be regarded as constants.

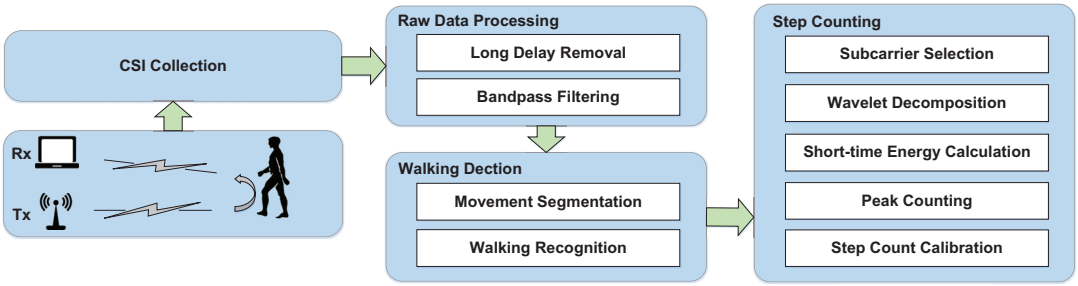


Fig. 2. System flow of WiStep.

4 WISTEP OVERVIEW

WiStep is devised to count walking steps using pervasive WiFi signals and it doesn't require users to carry or wear any special device. However, there exist two challenges:

(1) Walking recognition is the first step for step counting, it is challenging for WiStep to detect the starting points and endpoints of movements with the noisy CSI measurement, and further recognize the walking movement from various daily activities, *e.g.*, sitting down and waving a hand.
(2) Apart from movements of feet or legs, it always involves many other movements like breathing, swinging arms and swaying when a person is walking, which deliberately blurs the impacts of feet and legs on CSI values.

To address these challenges, we propose the processing procedures as depicted in Fig. 2.

In WiStep, the transmitter (a wireless router) is equipped with one directional antenna, and the receiver (a desktop) has 3 omni-directional antennas. While the system is working, the transmitter continuously sends wireless packets to the receiver, and the receiver collects the packets and extracts CSI measurements at a rate of 1000 Hz.

For Raw Data Processing, we first convert the frequency-domain CSI to time-domain Channel Impulse Response (CIR) by Inverse Fast Fourier Transform (IFFT) and remove the signals with long propagation delays. Then we utilize a Butterworth bandpass filter to wipe out the interference stemming from the high-frequency noise and some low-frequency body movements, such as breathing.

For Walking Detection, we refer to our previous work [20] and apply a dynamically self-adapting threshold to detect the start and end points of movements. And we apply a simple but efficient metric called Weighted Energy of Interest (WEI), which focuses on the energy of frequency components of walking, rather than some machine learning methods to recognize the walking movement.

For Step Counting, we first dynamically select the subcarriers with largest variances in each CSI stream and discard the subcarriers which are severely distorted by noise. Then we conduct wavelet decomposition on the selected subcarriers to compress the data and utilize the coefficients which represent the frequencies of feet and legs to construct a step profile. Short-time energy of the step profile is calculated to analyze and get the step count. Finally, based on the characteristics of Multiple-Input Multiple-Output (MIMO) and OFDM, we execute the intra-stream (averaging the step counts derived from multiple subcarrier of a single CSI stream) and inter-stream calibrations (averaging the step counts among multiple CSI streams) to optimize the final step count.

In what follows, the detailed processing procedures will be introduced.

5 CSI COLLECTION

Since the 2.4 GHz band of IEEE 802.11n standard is narrower and more crowded than the 5 GHz band, the latter is a much better choice for less inter-channel interference and more reliable communication. Moreover, the wavelength of 5 GHz band is shorter, which leads to better movement speed resolution. Therefore, we run WiStep on the 149# 802.11n channel (5 GHz band). Specifically, the central frequency of the channel is 5.745 GHz without the dynamic frequency selection (DFS) and transmit power control (TPC) constraints of Federal Communications Commission (FFC), and the signal wavelength is 5.22 cm. A commercial WiFi router along with a laptop serves as the transmitter, and the router is equipped with a directional antenna which can help to enhance the signals in its coverage and reduce the interference coming outside the coverage. A desktop mounted with an Intel 5300 NIC (equipped with 3 omni-directional antennas) serves as the receiver, and the firmware of the NIC is modified using the tool in [13] so as to report CSI measurements.

WiStep employs a fixed-size receive window to collect the CSI measurements. A set consisting of full CSI measurements in the receive window is regarded as the minimum processing unit. Considering human activities in traditional indoor environment introduce frequencies of no more than 300 Hz in CSI amplitudes [39], thus WiStep is configured with a sampling rate of 1000 Hz that is larger than the Nyquist sampling rate. WiStep organizes the CSI values of each CSI stream to form a matrix $\mathbf{C}_i(i = 1 \cdots N_C)$ with dimensions $N_S \times L$, where $N_C = 3$ and $N_S = 30$ separately denote the number of streams and subcarriers in each stream when there are 3 Rx antennas and 1 Tx antenna. L is the length of the receive window. We only utilize the CSI amplitude and ignore the CSI phase to eliminate the impact of Carrier Frequency Offset (CFO) as suggested in [39]. Next, we focus on one stream to facilitate depicting the downstream processing procedures.

6 RAW DATA PROCESSING

In this section, we describe some details of two preprocessing techniques, called long delay removal and bandpass filtering, to mitigate the interference from the complex indoor environment.

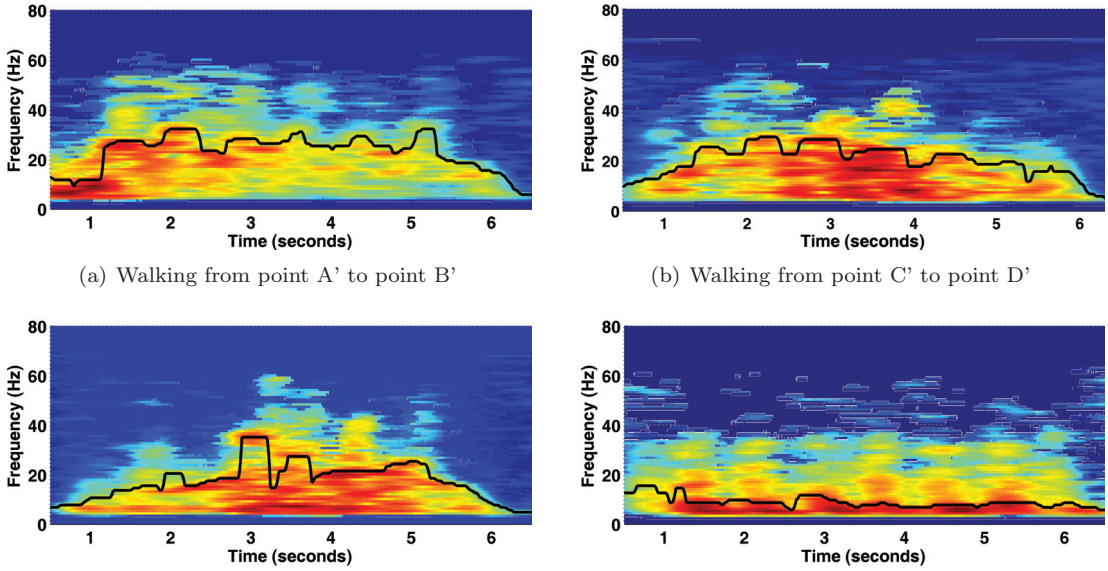
6.1 Long Delay Removal

According to the above multipath propagation model, the received signals consist of multiple copies of the signals with different propagation delays. This phenomenon can also be characterized by Channel Impulse Response (CIR) which is the inverse Fourier transformation of CFR. Theoretically, we can separate every signal with certain propagation delay from CIR, but the bandwidth of the working channel is 20 MHz, and the time resolution of CIR is approximately $1/20\text{MHz} = 50$ ns [41]. Therefore, we can't distinguish every signal from the low-resolution CIR but $N_S(= 30)$ signal clusters with discrete time delays.

The signals with long propagation delay probably are reflected by some static or dynamic objects which are far away from the transceivers, and these signals are useless and can distort the CSI amplitudes for WiStep. Besides, previous study shows the maximum delay in general indoor environment is less than 500 ns [16]. Thus, we transform each CSI measurement into time-domain CIR by IFFT and remove the components whose propagation delay is longer than 500 ns, which means the last 20 components of CIR are set to be 0. Then we convert the processed CIR back to CSI, using Fast Fourier Transformation (FFT).

6.2 Bandpass Filtering

Despite of removing multipaths with long propagation delay, the raw CSI measurements reported by commodity WiFi devices contain significant high-frequency noise due to the propagation channel and/or



(a) Walking from point A' to point B'

(b) Walking from point C' to point D'

(c) Walking from point A to point D

(d) In-place walking at point O

Fig. 3. Spectrograms of walking along the different routes.

hardware imperfection, *e.g.*, power control uncertainty [41], and low-frequency interference caused by slight body movements and/or environment changes, *e.g.*, chest movement while breathing. To better illustrate which frequency range should be reserved for walking recognition and step counting, we ask a volunteer to perform four walking movements in our laboratory as shown in Fig. 9(a), *i.e.*, (i) walking along the dash line from point A' to point B'; (ii) walking along the dash line from point C' to point D'; (iii) walking along the solid line from point A to point D; (iv) in-place walking at point O. We plot the spectrograms of the four walking instances in Fig. 3. The spectrograms and the black lines are generated using the methods proposed in WifiU [38], where a sliding window is first used to split the time-series CSI amplitudes into many pieces, then the FFT is performed on each CSI piece to get the time-frequency spectrogram, and finally the black lines which roughly represent the volunteer's torso frequency are estimated using the percentile method. Since the deployments of WiStep are different from WifiU, the derived spectrograms look a little different from WifiU's, but we can still get some important information:

(1) The hot color represents high FFT amplitude and generally corresponds to the torso reflection (torso has larger reflection area than other body parts), while the cold colors like cyan in the high frequency range of 40∼60 Hz correspond to the reflections of arms/feet/legs (with smaller reflection areas).

(2) The frequency distribution range in the spectrogram of in-place walking is narrower than the other three inter-point walking, since the moving distances of feet and legs for the in-place walking are shorter than the other three movements, which leads to lower moving speeds as well as lower frequencies (30∼40 Hz) in CSI values. Besides, for in-place walking, the average estimated torso frequency is around 10 Hz that is equivalent to 0.26 m/s (=10Hz×0.0522m/2), this probably is the speed of upper legs since the torso speed of the volunteer is almost null.

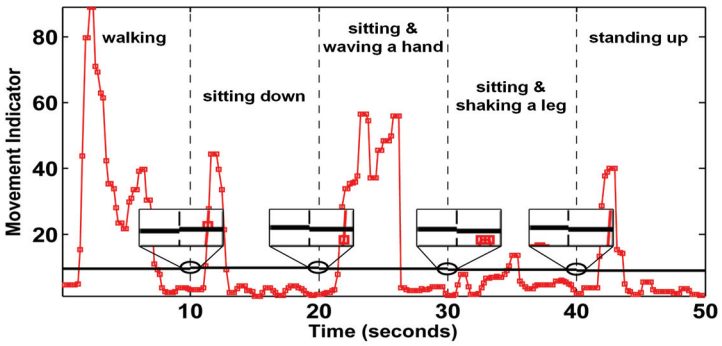


Fig. 4. The result of movement segmentation.

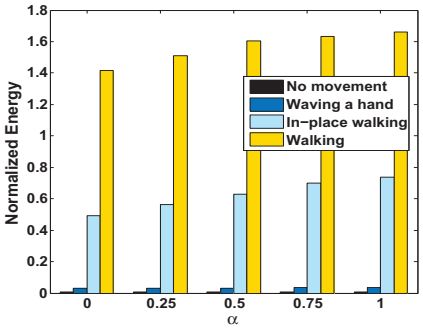


Fig. 5. The normalized WEI varying with the change of balance factor α.

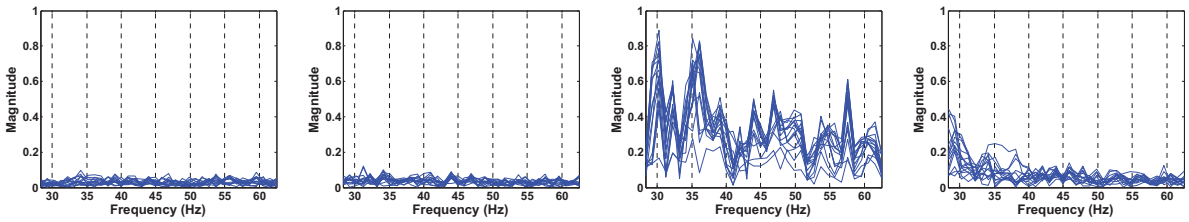
(3) The spectrograms of the first three walking movements with different walking directions have similar frequency distributions, and the averages of the black lines within the steady phases are around 26∼29 Hz, which are equivalent to the speeds of 0.68∼0.76 m/s.

Therefore, considering that the frequency components induced by walking are approximately between 20 Hz and 60 Hz in CSI amplitudes, WiStep exploits the Butterworth bandpass filter, which guarantees high fidelity of reserved signals in the passband, to eliminate the high-frequency and low-frequency noise, and the upper and lower cutoff frequencies of the Butterworth filter are set as 80 Hz and 5 Hz, respectively. The lower cutoff frequency is determined by the trade-off between fully eliminating low-frequency interference and the accuracy of step counting. The direct current component (0 Hz) of each subcarrier is also filtered by the bandpass filter. After removing the noise, WiStep introduces weighted moving average to further denoise and smooth the CSI amplitudes.

7 WALKING DETECTION

7.1 Movement Segmentation

Before recognizing walking movement from various daily activities, we first need to detect the start and end points of each movement in the receive window. We refer to our previous work [20] and modify the gesture detection method to obtain a satisfactory segmentation result. Specifically, we first cut the CSI stream into small bins using a 250-point sliding window (its length is a quarter of the sampling rate, and the sliding step is 200), and arrange each bin in columns to form a matrix $\mathbf{M}$ with dimensions 250×30. Then we conduct PCA on each matrix $\mathbf{M}$: (i) calculating the correlation matrix $\mathbf{M}^T \times \mathbf{M}$; (ii) performing the eigendecomposition of the correlation matrix to obtain the eigenvectors $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k$ (k is the number of leading eigenvectors to be learned, $1 \leq k \leq 30$); (iii) constructing the principal components $\mathbf{h}_i = \mathbf{M} \times \mathbf{q}_i$ $(1 \leq i \leq k)$. Subsequently, we define a movement indicator $\mathcal{I} = \sigma^2_{\mathbf{h}_2}/\Delta_{\mathbf{q}_2}$ to indicate the occurrence of movements, where $\sigma^2_{\mathbf{h}_2}$ denotes the variance of $\mathbf{h}_2$ and $\Delta_{\mathbf{q}_2}$ denotes the mean of the first difference of $\mathbf{q}_2$. When a movement happens, the CSI subcarriers become correlated and fluctuate obviously, thus $\Delta_{\mathbf{q}_2}$ becomes smaller and $\sigma^2_{\mathbf{h}_2}$ becomes larger, which results in a larger $\mathcal{I}$ and vice versa. Selecting the second component $\mathbf{h}_2$ is a trade-off between noise interference and computational complexity. In addition, a 5-point median filter is applied to smooth the calculated results to facilitate the segmentation of movements. Finally, we detect the start and end points of movements based on the noise level $L_{noise}(t)$, which is derived using the Exponentially Weighted Average method, where t denotes



(a) *No movement* (b) *Sitting and waving a hand* (c) *Walking* (d) *In-place walking*

Fig. 6. The normalized magnitudes of FFT coefficients for different movements.

the cycle index of CSI collection. The noise level $L_{noise}(0)$ is initialized with the mean of $\mathcal{I}$'s of the CSI bins when there is no activity in the receive window. Then $L_{noise}(t+1)$ is dynamically updated by the following rule during the working period:

$$L_{noise}(t+1) = (1-\gamma) \times L_{noise}(t) + \gamma \times \sigma^2_{\mathcal{I}}(t), \tag{3}$$

where the coefficient γ is set to 0.1 as suggested in [38], and $\sigma^2_{\mathcal{I}}$ is the variance of $\mathcal{I}$. We set the threshold as $3 \times L_{noise}(t)$ for current data receiving cycle t to detect the start and end points of movements in the receive window. The movement indicators below the threshold are employed to calculate $\sigma^2_{\mathcal{I}}(t)$, and only when $\sigma^2_{\mathcal{I}}(t) > 0$, the update process is triggered to update the noise level for the next cycle.

A segmentation example is illustrated in Fig. 4, where the length of the receive window is set as 10 seconds. During a 50-second period, we ask the volunteer to perform five different movements: (1) *walking*, (2) *sitting down*, (3) *sitting and waving a hand*, (4) *sitting and shaking a leg*, and (5) *standing up*. The horizontal lines represent the dynamically self-adapting detection thresholds. Since the noise level changes slowly over time, the threshold lines almost coincide. The pieces above the threshold lines are regarded as movements. We observe that movement (4) has little influence on the signals and is basically covered by the background noise. In addition, we assume the period of a normal walking movement is at least 3 seconds, which is reasonable in practice. Thus, the short-time movements such as *sitting down*, *standing up* in Fig. 4 are directly discarded without recognition.

7.2 Walking Recognition

Different from other movement recognition systems which are designed to identify various human activities [5, 39, 40], WiStep is specialized to recognize *walking* and count walking steps. Considering that different human activities can be identified using the frequency-domain motion energy, *i.e.*, the sum of squared magnitudes of FFT coefficients over the 0.1∼20Hz frequency passband [22], and according to our analysis in §6.2, we decide to exploit the frequencies within 30∼60 Hz, in which the frequencies are dominated by feet or legs, to recognize walking from other movements. Fig. 6 shows the normalized magnitudes of frequencies between 30 and 60 Hz for *no movement*, *sitting and waving a hand*, *walking* and *in-place walking*, respectively, over 15 subcarriers. We find that (i) among the subcarriers in one stream, the magnitudes of different frequency components stay relatively stable, which means we can reduce the computational resources by just selecting a single or a fraction of subcarrier(s) for walking recognition; (ii) the magnitudes of frequency components within 40∼60 Hz of *walking* are dramatically larger than the other three movements; (iii) for *in-place walking*, the magnitudes of frequencies between 30 and 40 Hz are larger than the magnitudes of frequencies between 40 and 60 Hz and other non-walking movements.

According to the observations, we propose anthor simple but efficient metric, called Weighted Energy of Interest (WEI) to recognize walking. WEI can be calculated as follows.

$$WEI = \frac{\alpha}{N_{F_{<40}}} \sum_{k=1}^{N_{F_{<40}}} m_k^2 + \frac{1-\alpha}{N_F - N_{F_{<40}}} \sum_{l=N_{F_{<40}}+1}^{N_F} m_l^2, \tag{4}$$

where $N_{F_{<40}}$ and N_F are the number of frequency components in the range of 30∼40 Hz and the components between 30 and 60 Hz, respectively. m_k is the magnitude of frequency component $k(k \in [1, N_{F_{<40}}])$, m_l is the magnitude of frequency component $l(l \in (N_{F_{<40}}, N_F])$, and α is the balance factor. Fig. 5 illustrates an example of how WEI varies with the change of α based on the FFT coefficients displayed in Fig. 6. By leveraging a suitable threshold, we can easily recognize the walking movements (including the in-place walking). Tuning α can enlarge the gap between in-place walking and other non-walking movements, which can help to improve the recognition accuracy. In this paper, α is set as 0.75 to both achieve a satisfactory step counting accuracy and reduce the influence of some low-frequency environment changes. The detailed performance will be discussed in §9.

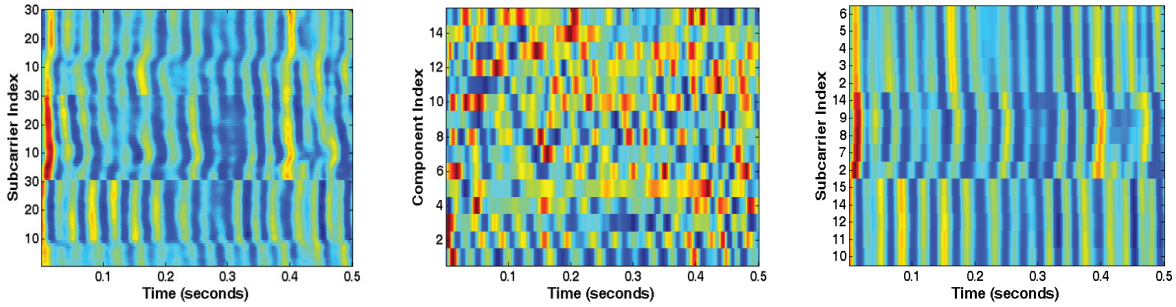
8 STEP COUNTING

For step counting, five detailed processing methods are described as follows.

8.1 Subcarrier Selection

In term of OFDM technology, different subcarriers have different carrier frequencies which result in different channel response according to Eq. (2). Due to frequency-selective fading, subcarriers are observed to have different sensitivities to human movements. Fig. 7(a) illustrates the intra- and inter-correlations of CSI streams for a 0.5-second walking fragment. Different colors in the heat map represent different values of CSI amplitudes, hot colors (*e.g.*, red) means large values and cold colors (*e.g.*, blue) means small values. At a certain time, different subcarriers in the same stream have different colors, which demonstrates they have different responses towards the movement, and the change is continuous. However, the change becomes discontinuous among different streams as there is a obvious horizontal boundary between the stream 1# and 2#. For a specific subcarrier, one alternating change in color roughly represents one wavelength change in the propagation path, and there are around 16 alternating changes in the 0.5-second period, which means the change frequency approximately is 32 Hz. Thus the walking speed is around $32 \times 0.0522/2 \approx 0.8$ m/s when the 149# WiFi channel (the central frequency is 5.745 GHz) is used in WiStep.

If we directly apply PCA on the 90 subcarriers and keep the first several principal components as suggested in [38], we can get the result like that in Fig. 7(b) where top 15 principal components are plotted. It is shown that the first two principal components have clear periodic characteristics, but the change patterns of the rest become indiscernible. Thus, only 2 data flows can be used to calibrate the final result. Considering that subcarriers with larger variances usually possess larger dynamic responses [27], thus we decide to dynamically select N_ν subcarriers ($1 \leq N_\nu \leq N_S$, N_S is the total number of subcarriers in CSI) with largest variances in each stream for step counting, which helps to reduce the negative influence of some disappointing subcarriers. Fig. 7(c) shows the concatenation of 15 selected CSI subcarriers, and every 5 adjacent subcarriers have the largest variances in the corresponding stream. All the 15 subcarriers show similar and clear change patterns, and the step counts derived from these subcarriers can be combined for a good result.



(a) (b) (c)

Fig. 7. Correlations of CSI subcarriers among different CSI streams and correlation of PCA components. (a) Bottom 30 subcarriers belong to stream 1#, middle 30 subcarriers belong to stream 2# and top 30 subcarriers belong to stream 3#. (b) Top 15 PCA components derived from the combination of 90 CSI subcarriers plotted in (a). (c) The concatenation of 15 selected CSI subcarriers, every 5 adjacent subcarriers have the maximum variances in the corresponding stream.

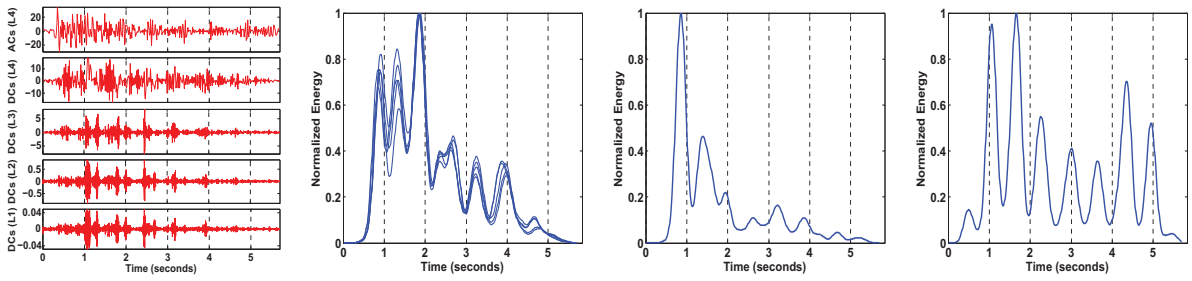
8.2 Wavelet Decomposition

Discrete Wavelet Transform (DWT) enables the ability of multi-resolution analysis of discrete signals in both time and frequency domain, and it provides high time resolution for movements with high frequencies and high frequency resolution for movements with low speeds. Generally, in the process of wavelet decomposition, the original signal first passes through a low-pass filter and a high-pass filter simultaneously and is decomposed into detail coefficients (DCs) and approximation coefficients (ACs). Then the ACs are set as the input of the same wavelet decomposition operation and decomposed in a recursive manner.

For each selected subcarriers of a *walking* segment, we perform a 4-level DWT decomposition with Daubechies db4 wavelet and scaling filters. Since the sampling rate of WiStep is 1000 Hz and the frequency ranges of adjacent DWT levels decrease exponentially, the frequency components 30∼60 Hz, which represent the speed of feet and legs, are mainly reserved in the DCs of the 4^{th} layer. The DWT result of the most sensitive subcarrier of the walking movement is illustrated in Fig. 8(a), and we observe that DCs with largest absolute value correspond to the fastest speed of a swinging leg in each step cycle, which is an important evidence for step counting. Another reason to applying DWT is that it can help to compress the data and further improve the computing speed. After DWT, we assemble the 4^{th} layer DCs S_i $(1 \leq i \leq N_\nu)$ of each subcarrier to construct the *step profile* $\mathbf{S}$ $(= \{S_1, S_2, \cdots, S_{N_\nu}\}^T)$.

8.3 Short-time Energy Calculation

In digital speech processing field, Short-time Energy (STE) is an effective metric of Voice Activity Detection (VAD) for finding the endpoints of the speech and distinguishing it from background noise [12]. Inspired by this, we calculate STE of the step profile to find out the walking steps. Specifically, each S_i in the step profile is first split into many short-time frames using a window function. Let $S_{i,n}(t)$ denote the n^{th} frame, then the STE E_n of frame $S_{i,n}(t)$ can be calculated as $E_n = \sum_{t=0}^{T-1} S_{i,n}^2(t)$, where T is the length of the frame. Subsequently, the weighted moving average method is introduced to filter out some local noise. Fig. 8(b) shows the STE of the top 5 subcarriers with largest variances, and the peaks of the STE curves basically correspond to the locally maximum values of the 4^{th} layer DCs in Fig. 8(a).



(a) (b) (c) (d)

Fig. 8. (a) Wavelet decomposition of the most sensitive subcarrier of stream 1#. (b) STE of 5 most sensitive subcarriers in the step profile. (c) STE of the subcarrier with smallest variance. (d) STE of an in-place walking instance.

Therefore, we can count the peaks to further calculate the number of walking steps. Besides, the curves in Fig. 8(b) almost overlap together which implies that the impacts of walking on different subcarriers are coincident. Fig. 8(c) shows the STE of the subcarrier with smallest variance, and there exists a severe distortion after the 2nd second. This kind of subcarriers are discarded in WiStep to prevent them from affecting the result. Furthermore, we also plot the STE of an in-place walking instance (of which the true step count is 7) in Fig. 8(d), where the curve shows obvious fluctuations, this is because the in-place walking involves less interference from the steady torso.

8.4 Peak Counting

To count the peaks which correspond to steps, three parameters, *i.e.*, peak width, peak height and inter-peak distance, are introduced to pick out the satisfactory peaks, where inter-peak distance is defined as the interval of two adjacent peaks and the parameters are learned using the training data.

We first use the thresholds of peak width and peak height to remove the small peaks which may caused by other body parts or environment noise. After that, the threshold of inter-peak distance is employed to find out the peaks whose intervals are shorter than the minimum time interval of two continuous steps, and these peaks are called as candidate peaks. We can't withdraw all the candidate peaks directly, because some of them are true steps. Hence, we develop a Candidate Peaks Fusion algorithm (see Algorithm 1 for details) to refine the candidate peaks and reduce the interference of false peaks. Finally, we can obtain the purified peak counts of the subcarriers in the step profile.

8.5 Step Count Calibration

With the characteristics of MIMO and OFDM, we execute the intra-stream, *i.e.*, averaging the peak counts of the selected subcarriers in each CSI stream, and inter-stream calibrations to optimize the final step count, which guarantees a more reliable and satisfactory result:

$$\begin{aligned} R_{STEP} &= \frac{1}{N_c} \sum_{i}^{N_C} \sum_{j}^{N_\nu} \omega_{i,j} p_{i,j}, \\ \sum_{j}^{N_\nu} \omega_{*,j} &= 1, \end{aligned} \tag{5}$$

ALGORITHM 1: Candidate Peak Fusion

Require: the set of candidate peaks' abscissas: $\mathcal{X} = \{x_1, x_2, \cdots, x_n\}$ $(x_1 < x_2 < \cdots < x_n,\ n \geq 2)$;
the minimum inter-peak distance: d $(d > 0)$.
Ensure: the refined $\mathcal{X}$
$i = argmin(x_{i+1} - x_i)$;
$\delta_{min} = x_{i+1} - x_i$;
while $\delta_{min} < d$ **do**
  $x_i = \frac{x_{i+1}+x_i}{2}$;
  Delete x_{i+1} from $\mathcal{X}$;
  $n = n - 1$;
  if $n \leq 1$ **then**
    Break;
  end if
  $i = argmin(x_{i+1} - x_i)$;
  $\delta_{min} = x_{i+1} - x_i$;
end while
return $\mathcal{X}$;

where R_{STEP} denotes the final estimated step count, N_C and N_ν are the number of CSI streams and subcarriers respectively, $p_{i,j}$ is the peak count of the j^{th} subcarrier of stream i, and $\omega_{i,j}$ is the normalized weight which is determined by the variances of subcarriers.

9 IMPLEMENTATION AND EVALUATION

In this section, we conduct experiments in two different indoor scenarios to evaluate the performance of WiStep.

9.1 Experimental Setup

9.1.1 **Hardware Deployment**. We set a laptop connected with a TP-LINK TL-WDR4300 wireless router as the transmitter, and a 3.3 GHz Intel(R) Core(TM) i5 CPU 4 GB RAM desktop equipped with an Intel 5300 NIC as the receiver. The transmitter operates in IEEE 802.11n mode at 5.745 GHz and is installed with a DB-Link directional antenna. The receiver has 3 external omni-directional antennas and its firmware is modified using the tool in [13] so as to report the CSI measurements. The transmitter and the receiver are placed on the same side as suggested in [38]. We utilize *pktgen* module of Ubuntu 12.04 operation system to continuously send packets from the transmitter to the receiver. Meanwhile, the sampling rate can nearly reach 1000 *packets/s* at the receiver side. The size of the receive window is empirically set as 6000 (6-second period) based on the sampling rate.

9.1.2 **Experimental Process**. To evaluate the performance of WiStep, we invite 50 volunteers (37 men and 13 women) to carry on the walking experiments. Except for 3 teachers, all the other volunteers are graduate students. Different from the gait recognition experiments, volunteers' ages, heights and weights aren't essential for WiStep, thus we don't record these privacy information. We conduct the experiments in two different indoor scenarios:

(1) *Laboratory*: the size of our laboratory is 8.5m×7.0m and there are merely some office furniture as shown in Fig. 9(a). The transmitter and the receiver are separately placed on two chairs that are 2 meters apart. We let the directional antenna facing the walking area, which can help to generate

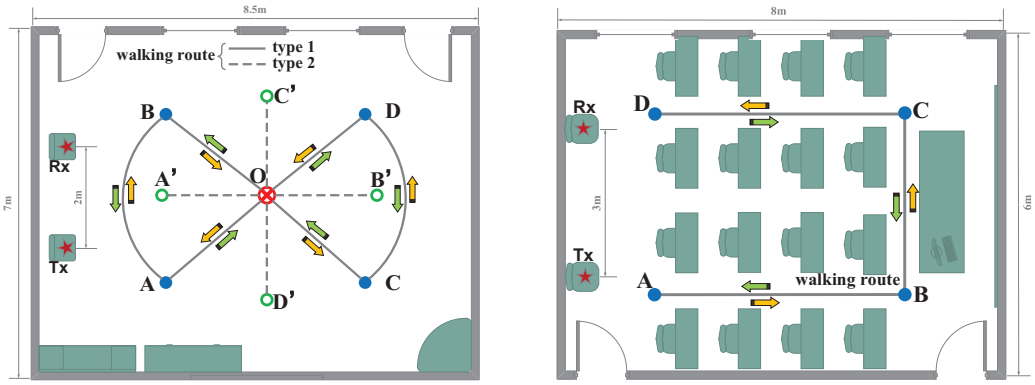


Fig. 9. Experimental scenarios: (a) Laboratory. (b) Classroom. Tx represents the transmitter and Rx represents the receiver. The solid and dash lines are the walking routes. The green and orange arrows indicate the walking directions.

more signals reflected off volunteers. Every volunteer is required to walk along the 8-shaped route (marked in solid line in Fig. 9(a)) for 40 rounds.

Concretely, in a single round, the volunteer first needs to perform an in-place walking movement at point O for at least 30 seconds, then he/she randomly selects one of the four directions to start walking. If the volunteer chooses the direction $O \rightarrow A$ or $O \rightarrow C$, he/she needs to follow the orange arrows and walk through the entire route back to point O; otherwise he/she needs to follow the green arrows. During walking, as long as he/she arrives at a marked point (except the point O), he/she needs to stay for 2 seconds without performing any movement, which can facilitate us aligning the data. Finally, again at point O, he/she performs the following 4 non-walking movements and each for 3 times: (1) sitting down (normal speed), (2) sitting and playing with a mobile phone (lasting for 10 seconds), (3) standing up (normal speed), (4) standing and waving a hand (lasting for 10 seconds).

(2) *Classroom*: the size of the classroom is 8.0m×6.0m, and it has a more complicated layout compared with the laboratory. As shown in Fig. 9(b), there are many desks and chairs which greatly enrich the propagation paths for the wireless signals and block some of the signals reflected off a person. The transmitter and the receiver are put on the chairs at a distance of 3 meters, and the transmitter is set to face the platform.

In the classroom, the volunteers also need to walk along the route for 40 rounds. Different from the above scenario, in each round, if the volunteer is at point A (D), he/she first performs an in-place walking (lasting for 30 seconds), then walks along the solid line and follows the orange arrows (green arrows) to point D (A). The volunteer also needs to stop at each marked point for 2 seconds while walking. Finally, he/she performs the non-walking movements as above.

The volunteers are asked to walk in their natural way and the CSI data are recorded by the receiver and processed offline. Besides, we also record the CSI data when there is no person in each scenario as "*no movement*" to initialize the noise level. The video cameras are applied to record the whole experimental process like that in [9, 25], and the ground truth is then counted by our recorders. All the video resources

are only used for the experiments and will be cleared for protecting the privacy of volunteers as soon as we finish the paper.

For the two scenarios, we randomly select the data of 35 volunteers to form the training set, and the data of the remaining 15 volunteers constitutes the test set. The training set is utilized to train the parameters of Walking Recognition and Peak Counting modules, and the test set is used to evaluate the performance of WiStep.

9.1.3 **Evaluation Metrics**. The following four metrics are employed to evaluate the performance of WiStep.

(1) *True Positive Rate (TPR)*: TPR is the ratio of the number of times that a true walking movement can be correctly identified as walking by WiStep.
(2) *False Positive Rate (FPR)*: FPR is the ratio of the number of times that WiStep incorrectly identify a non-walking movement as walking. TPR and FPR are mainly used to evaluate the Walking Detection module of WiStep.
(3) *Accuracy*: the accuracy of step counting for each volunteer is employed to evaluate the overall performance of WiStep, and it is defined as $1 - \frac{1}{N_r}\sum_{i=1}^{N_r}\frac{|R_e^i - R_t^i|}{R_t^i}$, where i is the round index, N_r represents the total rounds, R_e^i and R_t^i are the estimated step count and the ground truth value of the i^{th} round respectively. $R_e^i - R_t^i$ is regarded as the counting error.
(4) *Error Rate*: the error rate, which is also employed in [9], is defined as $\frac{R_e^i - R_t^i}{R_t^i} \times 100\%$, and it is used to evaluate the detailed step count error for each volunteer.

9.2 Walking Detection Evaluation

Recognizing a walking movement from various other daily movements is a basic and significant task for WiStep. With the method described in §7.1, WiStep can always realize an accurate segmentation result for detecting the start and end points of each daily movements. For walking recognition, we first cut the CSI data of the training set into a number of 6-second (which is equal to the length of the receive window) pieces and then label all the pieces with different movement types. Here we regard the in-place walking and walking along the route as the same type of movement, *i.e.*, "Walking", and all other movements as "Non-walking". Considering only the threshold of WEI needs to be learned, we can simplify the training process by calculating the WEI values of all the movement pieces and choose the value which can achieve highest recognition accuracy as the threshold.

At the test phase, we utilize the test data of 15 volunteers to simulate the real-time input CSI data flow for WiStep, and the recognition confusion matrix is shown in Fig. 10. From the matrix we can conclude that WiStep can correctly recognize the walking movement with overall TPR of 96.41% and FPR of 1.38%, which is comparable with the result (TPR of 98%) of Hidden Markov Model-based (HMM-based) recognition method proposed in [39], since we only utilize the simple threshold-based method with the metric WEI.

9.3 Step Counting Evaluation

For evaluating the step counting accuracy of WiStep, we first need to select all the data fragments that only involve walking movements to train the parameters of peak counting. Since all the volunteers are required to stop for 2 seconds at each point marked in Fig. 9(a) and Fig. 9(b), it's easy to separate all the useful training data. Totally, we obtain the training data of $(5+1) \times 40 \times 35 = 8400$ walking fragments in the laboratory scenario, and $(3+1) \times 40 \times 35 = 5600$ walking fragments in the classroom scenario.

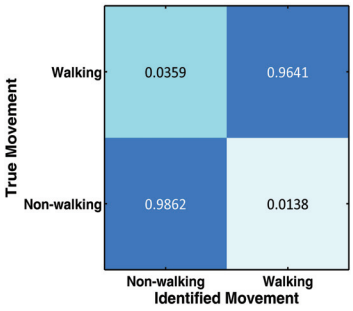


Fig. 10. The confusion matrix of walking recognition.

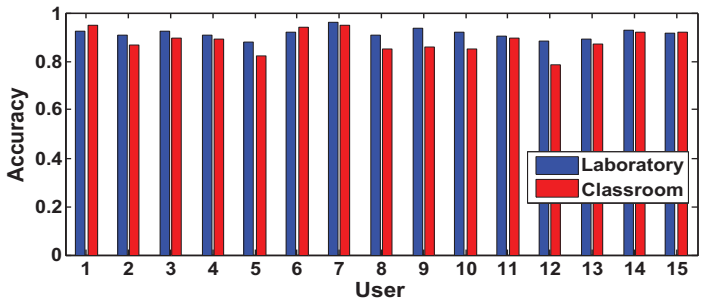


Fig. 11. The step counting accuracies of volunteer 1∼15 in different scenarios.

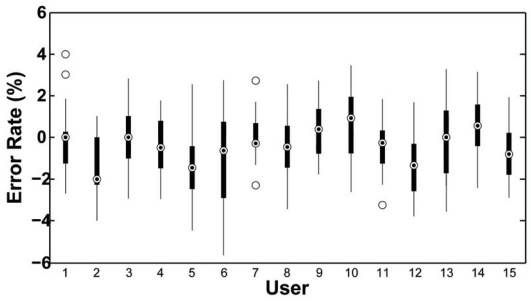


Fig. 12. The step counting error rates of volunteer 1∼15 in the laboratory.

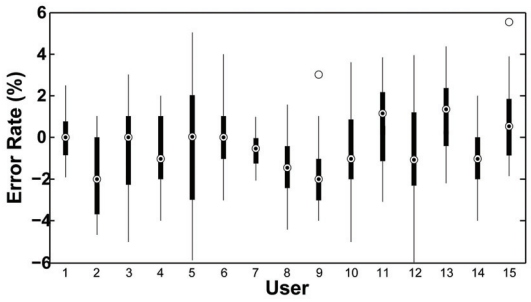


Fig. 13. The step counting error rates of volunteer 1∼15 in the classroom.

Besides, the data of "*no movement*" is also employed in the training phase to calculate the average peak height of noise.

To better test WiStep's performance of step counting, we only feed the walking fragments in the test set of each scenario to WiStep, *i.e.*, 3600 walking fragments collected from the laboratory and 2400 walking fragments collected from the classroom. Fig. 11 illustrates the step counting accuracies of volunteers $1 \sim 15$ in the two scenarios. As the figure shows, the step counting accuracies vary among different volunteers due to their walking habits. The overall average step counting accuracies for all volunteers in the laboratory and the classroom are separately 90.2% and 87.59%. The best step counting accuracies in the two scenarios belong to volunteer 7 (96%) and volunteer 1 (94.86%), respectively. Due to the complicated layout of the classroom, the existence of many desks and chairs may aggravate the multipath effect of WiFi signals, thus pulls down the step counting accuracy of WiStep slightly. However, the accuracy in the classroom scenario is still acceptable, and WiStep is robust to the changes of static reflectors.

In addition, Fig. 12 and Fig. 13 show the step counting error rates for the 15 volunteers in the two different scenarios and indicate that the majority of the step counting error rates lie in $[-4, 4]$.

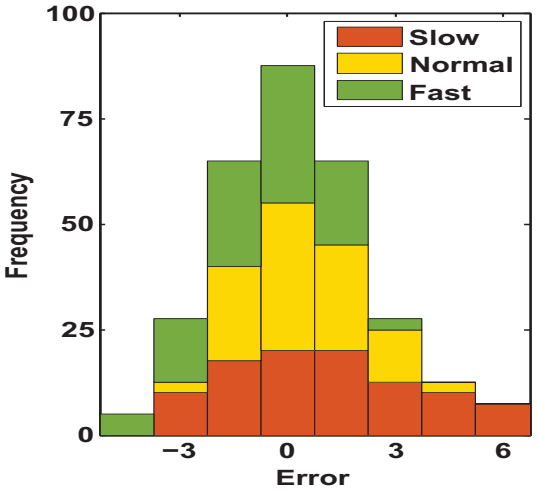


Fig. 14. The error distributions of step counting at different stride frequencies.

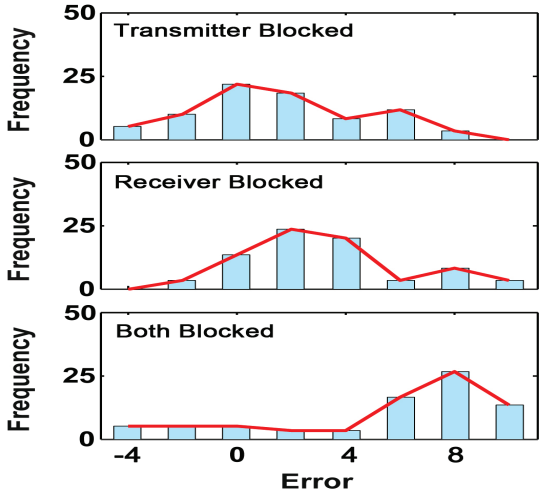


Fig. 15. The error distributions of step counting when the obstacle is placed at different positions.

10 DISCUSSION AND LIMITATIONS

10.1 Impact of Stride Frequency

Considering that different stride frequencies can lead to different change rates of CSI values, which may greatly influences the distribution of the frequency components. To evaluate the influence of stride frequency, we require volunteer 1 to walk at 3 different stride frequencies: slow ($\sim$50 *steps/minute*), normal ($\sim$75 *steps/minute*) and fast ($\sim$100 *steps/minute*, high enough for daily indoor walking). At each stride frequency, the volunteer is asked to walk for 100 rounds. Fig. 14 describes the error distributions of step counting at different stride frequencies. As the figure shows, the error distributions are similar to each other and the average error rates at slow, normal and fast stride frequency are 8%, 5% and 12% respectively.

10.2 Impact of Obstacle

The impact of obstacle is also taken into account in our system. We exploit an obstacle to obstruct the devices from facing the volunteer, and three representative conditions are considered: (i) the transmitter is blocked; (ii) the receiver is blocked; (iii) both the transmitter and the receiver are blocked. Fig. 15 separately shows the error distributions under the three conditions from top to bottom. We observe that when the transmitter or the receiver is blocked, the step counting errors mainly lie in $[-2, 2]$ and $[0, 4]$, and the accuracies are 93% and 85% respectively. However, when both the transmitter and the receiver are blocked, fewer signals can be emitted to or reflected from the volunteer, thus the step counting error increases dramatically which leads to a disappointing accuracy of about 38%.

10.3 Impact of Distance

Owing to the path loss and small-scale multipath fading, the signals get weaker with the increasing distance between the volunteer and the transmitter or receiver. To evaluate the impact of distance, volunteer 1 is asked to walk within three distance ranges: 0$\sim$2 m, 2$\sim$8 m and 8$\sim$12 m from the receiver.

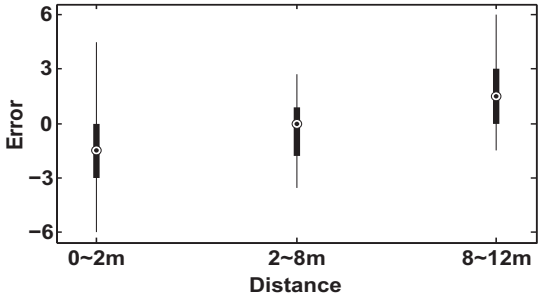


Fig. 16. The step counting errors for volunteers 1 when walking within different ranges.

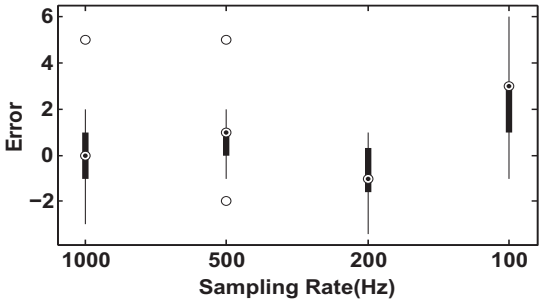


Fig. 17. The step counting errors for volunteers 1 with different sampling rates.

Table 1. The computational complexity of each processing procedure of WiStep.

Process	Complexity	Process	Complexity
Long Delay Removal	$O(LN_CN_S \log_2 N_S)$	Subcarrier Selection	$O(LN_CN_S)$
Bandpass Filtering	$O(LN_CN_S)$	Wavelet Decomposition	$O(4LN_CN_\nu)$
Movement Segmentation	$O(LN_S^2)$	Short-time Energy Calculation	$O(LN_CN_\nu)$
Walking Recognition	$O(LN_SF_s \log_2 F_s)$	Peak Counting & Calibration	$O(1)$

The step counting errors are illustrated in Fig. 16. It is observed that when volunteer 1 walks within the range of 2∼8 m, the error is smallest. The explanation is as follows: (i) When the volunteer walks in range 0∼2 m, the signals reflected by his arms or other body parts have high power and lead to strong interference. (ii)When the volunteer is within the range of 8∼12 m, the signals reflected by his feet or legs are weaken by the wireless channel and mixed with lots of chaotic noises. These all result in slight errors.

10.4 Impact of Sampling Rate

According to the Nyquist theorem, sampling rate has influence on the frequency-domain characteristics of CSI values. We conduct the evaluation of sampling rate by sampling the CSI data at different rates, such as 1000 Hz, 500 Hz, 200 Hz and 100 Hz. The result is shown in Fig. 17. When the sampling rate drops from 1000 Hz to 500 Hz, the step counting error almost keeps unchanged. However, when the sampling rate is set as 100 Hz, the error increases a lot. This is because some frequency components induced by legs are missed with the low sampling rate.

10.5 System Complexity Analysis

Table 1 illustrates the computational complexity of each processing procedure of WiStep when the receive window size is set to be L. In the table, $N_C = 3$, $N_S = 30$, $N_\nu = 5$ separately denote the number of CSI streams, subcarriers, and the top 5 most sensitive subcarriers in each CSI stream. F_s is the sampling rate. The top 2 most time-consuming procedures are Movement Segmentation (involves the PCA processing) and Walking Recognition (involves the F_s-point FFT processing over the time-series CSI amplitudes). The approximate computational complexity of WiStep is $O(LF_s \log_2 F_s)$, which enables WiStep to work in real time.

10.6 Limitations

WiStep is devised to count walking steps based on the multipath propagation model of the ubiquitous WiFi signals. However, there are some limitations:

(i) Like [1, 30], the directional antenna, which helps to enhance the signal strength and reduce the interference from its uncovered area or to other WiFi devices, is utilized in WiStep to transmit wireless signals. Although directional antennas aren't as ubiquitous as omnidirectional antennas for WiFi devices, we can deploy the directional antenna in the corner of a room so as to provide full coverage [30], which is equivalent to the effect of an omnidirectional antenna. WiStep can also work with omnidirectional antennas, but we leave the comparison between omnidirectional and directional antennas as future work.

(ii) Even though the directional antenna can help to reduce some external interference, if there are multiple people walking in the coverage at the same time, the signals reflected off different people are mixed together at the receiver side. Like many existing WiFi-based systems, for now, WiStep can't separate the mixed signals from multiple moving people. In our opinion, directly adding more WiFi devices can't solve the problem fundamentally, and an extra synchronization clock is probably helpful to calibrate the phase error and obtain more precise distance estimation.

(iii) We observe that WiStep can achieve high step counting accuracy when there are less obstacles between the human object and the transceivers, because the energy of the signals decays whenever they travel through the obstacles. This restrains WiStep from working in through-the-wall scenario.

(iv) During the experiments, the volunteers are required to walk normally (walking with their arms/hands swinging normally), and in order to fairly analyze the performance of WiStep among multiple volunteers, they are asked to walk in some predefined directions. However, we haven't conducted the comparison experiments like volunteers walking without swinging their arms/hands and/or walking in random directions. Since swinging arms/hands usually have the same cycle frequency as legs when people walking normally, and arms/hands induce much weaker reflected energy than that of torso/legs due to their smaller reflection areas, we think that arms/hands may have less influence on WiStep. Some specific analyses will be taken in our future work, and we will increase the generalizability of WiStep.

In future, we will also explore to introduce some synchronization mechanisms and more advanced signal processing techniques to improve WiStep's performance in more complex and fast-changing indoor environments.

11 CONCLUSION

In this paper, we designed a device-free step counting system, called WiStep, leveraging ubiquitous WiFi signals. Compared with many emerging radio-based human activity recognition applications, which mainly focus on classifying and recognizing daily activities, we payed more attention to the question "*to what extent can WiFi be utilized to count human walking steps*?" We explained the relation between different moving parts of a person and the variation of CSI values based on the multipath propagation model, and proposed to extract the frequency components induced by feet or legs rather than torso of the person for step counting, which enables WiStep to count the steps of in-place walking movement. We implemented WiStep using commodity WiFi devices in two different indoor scenarios, and various influence factors were taken into account to evaluate the performance of WiStep. The experimental results demonstrate that WiStep can realize overall step counting accuracies of 90.2% and 87.59% respectively in these two scenarios, and it is resilient to the changes of environment. WiStep is proven to be a reliable

and practical emerging step counting system and can be expanded to many correlative fields, such as indoor localization, gait recognition and exergames.

ACKNOWLEDGMENTS

The authors thank the reviewers for their insightful comments, and the volunteers for participating in our experiments. This work is supported by the National Natural Science Foundation of China (No. 61572456) and the Natural Science Foundation of Jiangsu Province of China (No. BK20151241).

REFERENCES

[1] Fadel Adib, Zachary Kabelac, and Dina Katabi. 2015. Multi-Person Localization via RF Body Reflections.. In *NSDI*. 279–292.
[2] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D Tracking via Body Radio Reflections.. In *NSDI*, Vol. 14. 317–329.
[3] Fadel Adib and Dina Katabi. 2013. See through Walls with Wi-Fi. In *ACM SIGCOMM*. Citeseer.
[4] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart Homes that Monitor Breathing and Heart Rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 837–846.
[5] Kamran Ali, Alex Xiao Liu, Wei Wang, and Muhammad Shahzad. 2015. Keystroke Recognition using WiFi Signals. In *Proceedings of ACM MobiCom*. 90–102.
[6] Moustafa Alzantot and Moustafa Youssef. 2012. UPTIME: Ubiquitous Pedestrian Tracking using Mobile Phones. In *IEEE WCNC*. IEEE, 3204–3209.
[7] Ferenc Aubeck, Carsten Isert, and Dominik Gusenbauer. 2011. Camera based Step Detection on Mobile Phones.. In *IPIN*. 1–7.
[8] Stéphane Beauregard. 2006. A Helmet-mounted Pedestrian Dead Reckoning System. In *2006 3rd International Forum on Applied Wearable Computing (IFAWC)*. VDE, 1–11.
[9] Agata Brajdic and Robert Harle. 2013. Walk Detection and Step Counting on Unconstrained Smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 225–234.
[10] Nicole A Capela, Edward D Lemaire, and Natalie C Baddour. 2014. A Smartphone Approach for the 2 and 6-minute Walk Test. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 958–961.
[11] R Delgado-Gonzalo, Celka, and *et al.* 2015. Physical Activity Profiling: Activity-specific Step Counting and Energy Expenditure Models using 3D Wrist Acceleration. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 8091–8094.
[12] Enqing Dong, Guizhong Liu, Yatong Zhou, and Yu Cai. 2002. Voice Activity Detection based on Short-time Energy and Noise Spectrum Adaptation. In *2002 6th International Conference on Signal Processing*, Vol. 1. IEEE, 464–467.
[13] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool Release: Gathering 802.11n Traces with Channel State Information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.
[14] Chunmei Han, Kaishun Wu, Yuxi Wang, and Lionel M Ni. 2014. WiFall: Device-free Fall Detection by Wireless Networks. In *Proceedings of IEEE INFOCOM*. 271–279.
[15] Chen-Yu Hsu, Yuchen Liu, Zachary Kabelac, Rumen Hristov, Dina Katabi, and Christine Liu. 2017. Extracting Gait Velocity and Stride Length from Surrounding Radio Signals. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2116–2126.
[16] Yunye Jin, Wee-Seng Soh, and Wai-Choong Wong. 2010. Indoor Localization with Channel Impulse Response based Fingerprint and Nonparametric Regression. *IEEE Transactions on Wireless Communications* 9, 3 (2010).
[17] Jeong Won Kim, Han Jin Jang, Dong-Hwan Hwang, and Chansik Park. 2004. A Step, Stride and Heading Determination for the Pedestrian Navigation System. *Journal of Global Positioning Systems* 3, 1-2 (2004), 273–279.
[18] Seung Young Kim and Gu-In Kwon. 2015. Interrupt-Based Step-Counting to Extend Battery Life in an Activity Monitor. *Journal of Sensors* 2016 (2015).
[19] Youngwook Kim and Hao Ling. 2009. Human Activity Classification based on Micro-Doppler Signatures using a Support Vector Machine. *IEEE Transactions on Geoscience and Remote Sensing* 47, 5 (2009), 1328–1337.
[20] Hong Li, Wei Yang, Jianxin Wang, Yang Xu, and Liusheng Huang. 2016. WiFinger: Talk to Your Smart Devices with Finger-grained Gesture. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 250–261.

[21] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-Sleep: Contactless Sleep Monitoring via WiFi Signals. In *IEEE RTSS*. IEEE, 346–355.
[22] Emmanuel Munguia Tapia. 2008. *Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure*. Ph.D. Dissertation. Massachusetts Institute of Technology.
[23] Nuria Oliver and Fernando Flores-Mangas. 2006. MPTrain: a Mobile, Music and Physiology-based Personal Trainer. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM, 21–28.
[24] Koray Ozcan and Senem Velipasalar. 2015. Robust and Reliable Step Counting by Mobile Phone Cameras. In *Proceedings of the 9th International Conference on Distributed Smart Cameras*. ACM, 164–169.
[25] Meng-Shiuan Pan and Hsueh-Wei Lin. 2015. A step counting algorithm for smartphone users: Design and implementation. *IEEE Sensors Journal* 15, 4 (2015), 2296–2305.
[26] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home Gesture Recognition using Wireless Signals. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*. ACM, 27–38.
[27] Kun Qian, Chenshu Wu, Zimu Zhou, Yue Zheng, Zheng Yang, and Yunhao Liu. 2017. Inferring Motion Direction using Commodity Wi-Fi for Interactive Exergames. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1961–1972.
[28] Cliff Randell, Chris Djiallis, and Henk Muller. 2003. Personal Position Measurement using Dead Reckoning. In *Proceedings of 7th IEEE International Symposium on Wearable Computers*. 166–173.
[29] Theodore S Rappaport et al. 1996. *Wireless Communications: Principles and Practice*. Vol. 2. Prentice Hall PTR New Jersey.
[30] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. 2015. WiBreathe: Estimating Respiration Rate using Wireless Signals in Natural Settings in the Home. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 131–139.
[31] Souvik Sen, Božidar Radunovic, Romit Roy Choudhury, and Tom Minka. 2012. You are Facing the Mona Lisa: Spot Localization using PHY Layer Information. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 183–196.
[32] Jungryul Seo, Yutsai Chiang, Teemu H Laine, and Adil M Khan. 2015. Step Counting on Smartphones using Advanced Zero-crossing and Linear Regression. In *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*. ACM, 106.
[33] Dave Tahmoush and Jerry Silvious. 2009. Radar Micro-doppler for Long Range Front-view Gait Recognition. In *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 1–6.
[34] David Tse and Pramod Viswanath. 2005. *Fundamentals of Wireless Communication*. Cambridge university press.
[35] Jeffrey J VanWormer. 2004. Pedometers and Brief E-counseling: Increasing Physical Activity for Overweight Adults. *Journal of Applied Behavior Analysis* 37, 3 (2004), 421–425.
[36] G. Wang, Y. Zou, Z. Zhou, k. wu, and L. Ni. 2014. We Can Hear You with Wi-Fi!. In *Proceedings of ACM MobiCom*. 593–604.
[37] Hao Wang, Daqing Zhang, Yasha Wang, Junyi Ma, Yuxiang Wang, and Shengjie Li. 2016. RT-Fall: A Real-time and Contactless Fall Detection System with Commodity WiFi Devices. *IEEE Transactions on Mobile Computing* (2016).
[38] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait Recognition using WiFi Signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 363–373.
[39] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and Modeling of WiFi Signal Based Human Activity Recognition. In *Proceedings of ACM MobiCom*. 65–76.
[40] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: Device-free Location-oriented Activity Identification using Fine-grained WiFi Signatures. In *Proceedings of ACM MobiCom*. 617–628.
[41] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise Power Delay Profiling with Commodity WiFi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 53–64.
[42] Zhenjie Yao, Zhipeng Zhang, and Liqun Xu. 2014. An Effective Algorithm to Detect Abnormal Step Counting based on One-Class SVM. In *2014 IEEE 17th International Conference on Computational Science and Engineering (CSE)*. IEEE, 964–969.
[43] Yunze Zeng, Parth H Pathak, and Prasant Mohapatra. 2016. WiWho: WiFi-based Person Identification in Smart Spaces. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks*. IEEE Press, 4.

Received May 2017; revised August 2017; accepted October 2017