On CSI and passive Wi-Fi radar for opportunistic physical activity recognition

Li, Wenda; Bocus, Mohammud Junaid; Tang, Chong; Piechocki, Robert J.; Woodbridge, Karl; Chetty, Kevin

Published in:
IEEE Transactions on Wireless Communications

DOI:
10.1109/twc.2021.3098526

Publication date:
2021

On CSI and Passive Wi-Fi Radar for Opportunistic Physical Activity Recognition

Wenda Li*, Mohammud J. Bocus †, Chong Tang*, Robert. J. Piechocki†, Karl Woodbridge‡, Kevin Chetty*
*Department of Security and Crime Science, University College London, UK
†Department of Electrical and Electronic Engineering, University of Bristol, UK
‡Department of Electronic and Electrical Engineering, University College London, UK
{wenda.li, chong.tang.18, k.woodbridge, k.chetty}@ucl.ac.uk, {junaid.bocus,r.j.piechocki}@bristol.ac.uk

***Abstract*—The use of Wi-Fi signals for human sensing has gained significant interest over the past decade. Such techniques provide affordable and reliable solutions for healthcare-focused events such as vital sign detection, prevention of falls and long-term monitoring of chronic diseases, among others. Currently, there are two major approaches for Wi-Fi sensing: (1) passive Wi-Fi radar (PWR) which uses well established techniques from bistatic radar, and channel state information (CSI) based wireless sensing (SENS) which exploits human-induced variations in the communication channel between a pair of transmitter and receiver. However, there has not been a comprehensive study to understand and compare the differences in terms of effectiveness and limitations in real-world deployment. In this paper, we present the fundamentals of the two systems with associated methodologies and signal processing. A thorough measurement campaign was carried out to evaluate the human activity detection performance of both systems. Experimental results show that SENS system provides better detection performance in a line-of-sight (LoS) condition, whereas PWR system performs better in a non-LoS (NLoS) setting. Furthermore, based on our findings, we recommend that future Wi-Fi sensing applications should leverage the advantages from both PWR and SENS systems.**

***Index Terms*—Passive Wi-Fi Radar, Channel State Information, Doppler, Wireless Sensing**

I. Introduction

With the fast growing ageing population, there are increasing concerns that health conditions such as cardiovascular diseases, mental health issues and diabetes will become more prevalent and increase the burden on national healthcare services. Hence there is a greater need than ever to provide efficient technology solutions for ambient assisted living and e-healthcare services [1]. Such systems are extremely helpful for improving quality of life and preventing health risks where early interventions are critical. Compared to other technologies used in healthcare monitoring such as cameras and wearable sensors, Wi-Fi based sensing technology is considered as an ideal solution because it does not produce images or identify people being monitored, which alleviates privacy concerns. Furthermore, it performs uncooperative sensing to overcome the limitations of wearable technology, such as low compliance rate, especially amongst the elderly, and discomfort for some users (e.g., those with skin conditions). Nowadays, Wi-Fi devices are readily available in almost all indoor environments, whether residential or commercial. Wi-Fi sensing systems have been studied for activity and gait recognition [2], fall detection [3], gesture recognition [4], and intrusion detection [5]. The main concept behind Wi-Fi sensing is that a moving person affects the wireless signal in terms of frequency shift, propagation paths and signal attenuation. Consequently, the communication channel becomes time-varying and this can be exploited for a number of sensing applications.

Wi-Fi sensing may be categorized into three types of systems: received signal strength (RSS) based, radar based and channel state information (CSI) based. Multiple parameters can be computed from received signals such as angle of arrival (AoA), time of flight (ToF), time difference of arrival (TDoA) and Doppler frequency shift (DFS). The purposes of these parameters are varied. For example, the coarse-grained RSS has been used for device-free indoor localization with fingerprinting (FP) method [6]. However, it suffers from temporal fluctuations in complex indoor environments because of multipath induced fading even in a static background. Also, RSS approaches require a substantial radio-map survey in the offline training phase and labour-intensive updates for changes in the background [7]. AoA is a useful parameter that is obtained by calculating the phase difference of the signals arriving at multiple antennas [8], where a large antenna array is normally required. ToF gives the relative distance between the transmitter and receiver by calculating the arrival time of the direct wave [9]. In a Wi-Fi system, using only the ToF parameter for localization is quite challenging since it is limited by the signal bandwidth. For instance, due to the low sampling rates of the 20 and 40 MHz channel bandwidths (50 ns and 25 ns time resolutions, respectively), the signal may arrive between sampled intervals, giving rise to distance estimation errors in the order of several meters.

Our previous work [10] has shown the potential of PWR system to detect small movements of the chest wall in signs-of-life detection with DFS parameter. This prototype has been further extended for micro-DFS based activity event classification [11]. On the other hand, Wi-Fi CSI which represents how wireless signals propagate from the transmitter to the receiver across multiple paths [5], has been used in many sensing applications like activity recognition [12] (DFS), device-free tracking [13] (ToF/AoA) and fall detection [3] (DFS).

In this paper, we focus on two major approaches for activity recognition; PWR system which generates the target's ToF/DFS parameters and SENS system which provides CSI data that can be converted into ToF/DFS parameters. However, due to the 20 MHz bandwidth in the 2.4 GHz band, ToF

TABLE I: Overview of Some Recent Wi-Fi Sensing Works

Reference	System	Signal Processing	Machine Learning	Application	Performance
[6]	Distributed Wi-Fi AP	FP, compressive sensing, cluster	N/A	indoor localization	90% error of 2.7 mover 26 APs
[14]	SENS	Wavelet-based denoising, multipath mitigation	PCI, subcarrier selection, DTW	finger gesture recognition	93% accuracy over 8 finger gestures
[15]	SENS	PCA, thresholding	DWT, HMM	activity recognition	96.5% accuracy over 9 activities
[2]	SENS	STFT, spectrogram superimposition	SVM (radial basis function (RBF) kernel)	activity recognition	average false acceptance rate and false rejection rate of 8.05% and 9.54%
[3]	SENS	interpolation, segmentation	8 empirical features, SVM	fall detection	91% of sensitivity and 92% of specificity
[16]	SENS	chest motion modeled as Fresnel zone	N/A	breathing detection	show ability in 1 m from different orientations
[17]	PWR	CAF, CLEAN	N/A	long distance detection	detect a moving person at 17 m away and TTW
[18]	PWR	ECA, CAF	N/A	outdoor detection	detect a running person and a moving car
[19]	PWR	CAF, CLEAN	N/A	finger gesture and activity recognition	feasibility demonstration
[10]	PWR	CAF, CLEAN, micro Doppler extraction	N/A	breathing detection	detection range up to 1 m at different orientations
[11]	PWR	CAF, CLEAN	HMM, K-means clustering	activity recognition	80% accuracy with unsupervised learning over 6 activities

does not provide sufficient resolution for this work, and thus only the DFS parameter is used. Two prototypes have been implemented to demonstrate the feasibility of each approach. Experimental data from these prototypes was collected and timestamped using a network time protocol (NTP) server for synchronization purposes. Detection performance was verified by six activities from five people with different setups. Although both SENS and PWR systems can provide meaningful results for human activity recognition, there are still many challenges that need to be solved for real-world deployment. Compared to previous SENS [4], [5], [20] and PWR [10], [18], [19] studies, this work makes the following contributions:

- To the best of the authors' knowledge, this is the first work to demonstrate the difference between SENS and PWR systems with experimental data collected from real-world.
- We have implemented two systems for SENS and PWR to facilitate a direct comparison. Thorough experiments are performed to investigate the difference between the two systems for human activity recognition, and identify their sensitivities in terms of system layout, coverage and accuracy.
- Based on our observation, we provide a discussion on SENS and PWR systems in terms of their advantages and limitations in future development and deployment. We also assess potential improvements for Wi-Fi based sensing.

This paper is organized as follows: Related works are presented in Section II. An overview of Wi-Fi sensing is given in Section III. Signal processing for the SENS system is described in Section IV. Signal processing for the PWR system is explained in Section V. System design and evaluation are presented in Section VI. Section VII discusses the feasibility and limitations of the two approaches. Finally, conclusions are drawn in Section VIII.

II. Related Work

In this section, we compare previous works for SENS and PWR systems. Generally, DFS parameter can be obtained from either raw Wi-Fi signal or CSI data. However, the way that the two systems extract this parameter is vastly different. A comparison of these works is shown in Table I.

A. *Wi-Fi SENS System*

For a Wi-Fi system with multiple-input multiple-output orthogonal frequency division multiplexing (MIMO OFDM) capability, the CSI is obtained as a 3D matrix, consisting of complex values which can be broken down into amplitude and phase [5]. CSI measurements, in time domain, capture the changes in wireless signal due to the latter's interaction with surrounding objects or human activities and the observed patterns can be used for various purposes. Different Wi-Fi sensing applications have specific requirements in terms of their signal processing techniques and classification/estimation algorithms.

For example, [6] presents a compressive sensing based FP method for indoor localization by using RSS parameter. However, the radio map is very time consuming to build and needs calibration when the environment changes, which limits its potential application in a residential environment. Authors of WiFinger [14] extract fine-grained features from CSI data by using principal component analysis (PCA) and employ dynamic time warping (DTW) as the classifier for finger gesture recognition. They claim to achieve an accuracy of 95% for 8 finger gestures compared to 76% using RSS. The idea of WiFinger is that subcarriers in an OFDM signal are highly sensitive to small movements in the physical environment and result in fluctuations in the CSI. This phenomenon has also been used in device-free activity recognition [2], [15]. For example, CARM [15] proposes a CSI-speed model to quantify the relationship between CSI variation and human movement speed. The frequency component is extracted from CSI measurements using discrete wavelet transform (DWT), and a hidden Markov model (HMM) is used to build the CSI-activity model to classify human activities. Another work WifiU [2] uses the short-time Fourier transform (STFT) technique to transform the CSI measurements into spectrograms. The torso speed and cycle time of each gait are calculated as features and support vector machine (SVM) is used as the classifier. This work claims to achieve an accuracy of 92% for a human walking at a distance of 14 m. Fall detection is another important area in Wi-Fi sensing. RT-fall [3] proposed a system that is able to detect human falls automatically and segment these falls from other activities. CSI has also been used in breathing detection [16] with the understanding of the Fresnel Zone between transmitter and receiver.

One of the major challenge for a SENS systems is that changes in the surrounding environment can significantly affect the communication channel. Several approaches have been used to eliminate the training phase in each new environment. For example, [20] computes different metrics from the CSI measurements such as mean, standard deviation, etc., for people counting applications that only require training within that specific environment. Wigest [4] uses the RSS parameter for gesture recognition by extracting the frequency component with wavelet transform, and no calibration is required. Similarly, [15] converts CSI into Doppler spectrograms using STFT, thus avoiding any calibration. A common approach in studies mentioned above is that they convert the RSS/CSI into the form of DFS parameter to avoid calibration in the dynamic environment.

B. Passive Wi-Fi Radar

Aside from SENS system, passive radar can be also used for Wi-Fi sensing which exploits Wi-Fi access points as illuminators of opportunity. Passive radar has a long history in airborne tracking and detection, but only over the last decade has it been used for personnel detection [17]. The underlying concept of passive radar is to exploit the signals from third-party transmitters, to measure the time difference between the signal arriving directly to a reference receiver and the signal arriving via reflection from the object through a synchronized surveillance channel.

Signal processing for PWR system is more straightforward than the SENS system. It uses a cross ambiguity function (CAF) to generate ToF (bistatic distance) and DFS (bistatic velocity) parameters. However, due to the limited bandwidth of Wi-Fi APs (20 or 40 MHz), only DFS parameter is mainly used for indoor scenario. The advantage of DFS parameter, as discussed above, is that no calibration for the surrounding environment is required. An early attempt of the PWR [17] shows the feasibility of using Wi-Fi signal to detect personnel at a stand-off distance under through-the-wall (TTW) condition. [18] investigates a PWR system for outdoor scenarios, and successfully detects both ToF and DFS parameter for a moving car and a running human. [19] built a prototype based on the software defined radio (SDR) platform with real-time ability, and showed potential for several applications such as activity and finger gesture recognition. PWR has also been used for breathing detection. For instance, in [10] we demonstrate that micro-DFS parameter can be obtained from the chest motion. PWR for indoor localization can be achieved by examining the DFS parameter from a moving object with at least two separated channels [21]. However, accumulated error in the DFS parameter affects the localization accuracy [21].

The strong direct signal from a Wi-Fi AP is a major source of interference for a PWR system. Both physical (e.g. angular antenna nulling), and post-processing techniques such as adaptive filtering can be used to remove the direct signal interference (DSI). Work [18] proposed an extensive cancellation algorithm (ECA) that subtracts the direct signal from reflected signals based on the least square technique. However, ECA has a high computational load making real-time processing infeasible. Another work [17] uses a modified 'CLEAN' algorithm to suppress the dominant peak due to the direct signal with a self-ambiguity function which is calculated by the reference channel only and shares a similar structure with CAF.

III. Overview of Wi-Fi Sensing

A. Signal Model

OFDM waveform has been widely used in many Wi-Fi standards such as IEEE 802.11 a/g/n/ac. In an OFDM system, the bandwidth is shared among multiple overlapping but orthogonal subcarriers and due to the small bandwidth, each subcarrier experiences only flat fading in a frequency-selective fading wireless channel. Let the transmitted OFDM signal be defined as:

$$x(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} a_n e^{j\frac{2\pi}{T}nt}, \tag{1}$$

where T is the OFDM symbol period, N is the number of subcarriers and a_n is the nth symbol in the constellation symbol sequence such as QPSK or QAM. The received signal $y(t)$ consists of both the direct signal and multipath reflections. The reflections from a stationary object or a moving person can be represented as a summation of the delayed and phase shifted transmitted signal. The received signal can be written as:

$$y(t) = \sum_p A_p e^{j2\pi f_d t} x(t-\tau) + n(t), \tag{2}$$

where p is the number of reflected paths, A_p, τ, f_d are the attenuation factor, delay and Doppler shift for the p-th path respectively, and $n(t)$ is the additive white Gaussian noise (AWGN).

In the frequency domain, the transmitted signal $X(f_c, t)$ and received signal $Y(f_c, t)$, with carrier frequency f_c are related through the expression $Y(f_c, t) = H(f_c, t) \times X(f_c, t)$, where $H(f_c, t)$ represents the channel frequency response (CFR) at carrier frequency f_c, measured at time t. $H(f_c, t)$ can be expressed as:

$$H(f_c, t) = e^{j2\pi\Delta f_c t} \sum_p A_p(f_c, t) e^{j2\pi f_d (t-\tau)}, \tag{3}$$

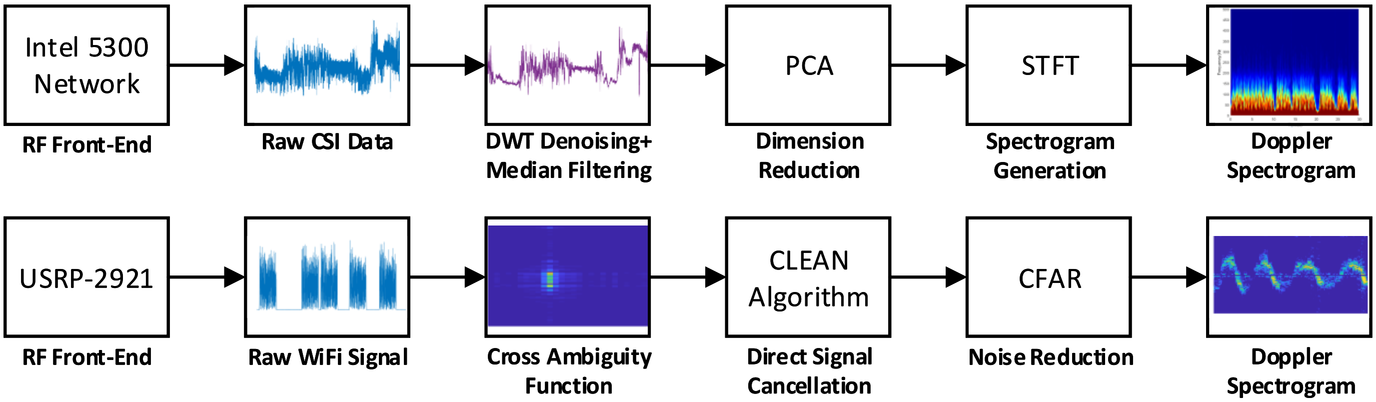

Fig. 1: Block diagram overview of SENS (top) and PWR systems (bottom)

where $e^{j2\pi f_d(t-\tau)}$ is the phase shift with f_d being the Doppler frequency and τ the propagation delay. $e^{j2\pi\Delta f_c t}$ is the phase difference between transmitter and receiver due to the sampling frequency offset (SFO) and sampling time offset (STO). Although the mechanism of SENS and PWR system is different, however, the key idea of both systems is to detect the changes in the communication channel caused by moving targets and at the same time remove interference from surrounding objects as well as the geometry of transmitter and target reflection.

B. System Model

Here we present an overview of the signal processing used in the PWR and SENS systems for human activity sensing. This is also summarized by the block diagram in Fig. 1. The SENS system is based around communication techniques with a commercial off-the-shelf (COTS) Intel 5300 [22] NIC as the RF front-end, which generates CSI measurements and is stored for off-line processing. The PWR system is based on the radar technique with a universal software radio peripheral (USRP) platform as the RF front-end, which samples the raw wireless signal and digitizes it for real-time processing on a desktop.

In order for a Wi-Fi receiver to decode correctly the transmitted signals in a wireless medium, the propagation characteristics of the channel must be known. In this regard, a training sequence that is known by both the transmitter and receiver is sent in each packet to obtain the channel estimate. This process is often referred to as channel sounding. The channel estimate is known as CSI and for a MIMO-OFDM system, it is a matrix consisting of complex values for each subcarrier. The equalizer uses the CSI to reverse the effects of the channel on the transmitted signal such as multipath propagation, attenuation, phase shift, etc. In the IEEE 802.11n standard, the training sequences are known as high throughput long training fields (HT-LTF) and they are sent as part of the preamble for the receiver to obtain the CSI [23]. On the other hand, PWR system (with radar technique) correlates the transmitted signal $x(t)$ and received signal $y(t)$ [24] to detect the Doppler shift (DFS) f_d and propagation delay (ToF) τ. PWR follows the structure of passive radar system, that is, it has a 'reference channel' to recover the transmitted signal and

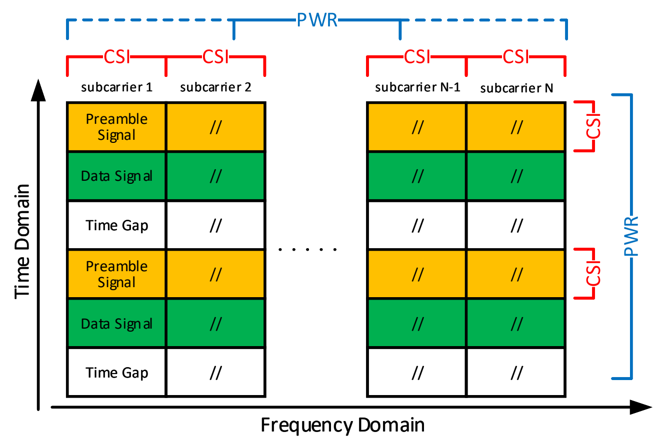

Fig. 2: Demonstration of SENS and PWR mechanisms in time and frequency domains.

several 'surveillance channels' to capture the reflected signals from different angles to provide spatial information.

As discussed before, only the DFS parameter will be used in this work. Thus we output the Doppler spectrogram from the two systems as it can present meaningful information about the activity and it is also insensitive to static objects in the background. There are three major stages in the SENS system: (1) DWT denoising & median filter to filter out in-band noise and preserve the high frequency components; (2) CSI data size reduction by using PCA and just keep meaningful PCA values and (3) convert these values into Doppler spectrogram using the STFT method. There are also three major stages in the PWR system: (1) a Cross Ambiguity Function (CAF) is used to generate a range-Doppler surface based on the transmitted and reflected signals; (2) a CLEAN algorithm [17] is needed to suppress the direct signal from signal source which is the major interference for a PWR system and (3) noise reduction is done by applying a Constant False Alarm Rate (CFAR).

Afterwards, these Doppler spectrograms will be used to train a deep neural network and calculate accuracy for each system. More details on the signal processing of each system are given in sections IV and V.

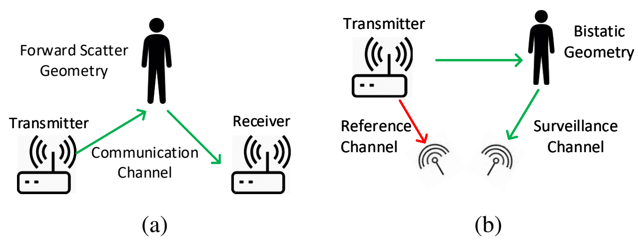

Fig. 3: Layout: (a) SENS system and (b) PWR system

C. Mechanism

The mechanisms of the SENS and PWR systems in time and frequency domains are illustrated in Fig. 2. CSI measurement characterizes how wireless signals propagate from the transmitter to receiver based on the preamble in a Wi-Fi packet. The pre-defined sequence is used to generate the corresponding CSI measurement. However, SENS system ignores the data signal and hence does not take full advantage of a whole Wi-Fi packet. In comparison, the PWR system does not require details of the preamble or data in a Wi-Fi packet. To ensure Doppler sensitivity and signal content, PWR records signals for a longer duration than CSI. The advantage of the PWR system is that it can use both the preamble and data signal, whereas it considers the time gap between two packets to be redundant noise. The activity detection performance of both the SENS and PWR systems depends on the frequency of the received Wi-Fi packets, where the typical default setting of a commercial Wi-Fi AP (10 beacon frames per second) is not sufficient for sensing.

SENS systems make estimates about the communication channel at each subcarrier (in frequency domain). These measurements provide fine-grained features but they normally result in a considerable size for sensing applications. On the other hand, PWR system does not process the OFDM signal on subcarrier basis but treats each OFDM symbol as one signal. For this reason, PWR system cannot access the information within each subcarrier. The bandwidth of the PWR system is adjustable from the full Wi-Fi spectrum to a single tone. It is true that the channel information from each subcarrier provides better resolution than the PWR system which processes the Doppler shift from the whole signal instead of individual subcarriers. However, the variations across all subcarriers may not be vastly different.

SENS and PWR have different configurations as shown in Figure 3. The major difference is that PWR system has two channels (surveillance and reference channels) compared to a single channel (communication channel) in the SENS system. The reference channel aims to recreate the originally transmitted signal as needed in a passive radar system to compute the range/Doppler information with the surveillance channel [25]. Usually the reference channel is set up by pointing a directional antenna towards the signal source to capture the transmitted signal. However, the reference channel is unstable in reality since it may not perfectly recreate the signal due to multiple environmental factors. Some minor interference in the reference channel will not significantly affect the PWR's performance. In comparison, the predefined preamble signal in the SENS system is more reliable to generate the CSI measurements. Additionally, this also simplifies SENS to a single channel system (No reference or surveillance channels required).

SENS system captures the variations in the communication channel between the transmitter and receiver due to motion. It works best, i.e., it is most sensitive to the variations caused by human activity when there is a LoS path between the transmitter and receiver. This geometry is also known as the forward scatter, where the angle between transmitter-target and target-receiver is around 180 degrees [25]. However, forward scatter is not ideal for PWR systems since the Doppler information is lost in a LoS setting [26], which therefore reduces the system's sensitivity. As a passive radar system, PWR performs better in a monostatic geometry, where the angle between transmitter-target and target-receiver is smaller than 90 degrees [25]. Since the two systems work best in different physical geometries, they have different coverage sensitivities and their Doppler spectrograms will differ for the same geometry.

IV. Signal Processing for SENS System

A. SFO and STO Removal (Phase Calibration)

In a practical Wi-Fi system, raw CSI measurements are affected by phase offsets as the hardware and software are not ideal. STO and SFO are caused by non-synchronized sampling clocks and frequencies between the transmitter and receiver, respectively. Since in this work we focus on human activity sensing, the time-domain CSI amplitude variations are enough for this purpose as they exhibit different patterns for different activities. Nonetheless, the phase is calibrated as in [27] where a linear transformation is applied to the raw phase data to eliminate the phase offset. The measured phase $\hat{\phi}_i$ of the ith subcarrier be expressed as:

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N}\delta t + \beta + Z, \qquad (4)$$

where ϕ is the true phase, β is the phase offset due to carrier frequency offset, δt is the timing offset between the transmitter and receiver, k_i is the index of the ith subcarrier and Z is the measured noise. In the Intel 5300 CSI tool [28], $i \in \{1, 30\}$ and N is the fast Fourier transform (FFT) size. For example, $N = 64$ for a 20 MHz Wi-Fi channel in IEEE 802.11 a/g/n. The terms δt, β and Z make it difficult to obtain the true phase from Wi-Fi NICs. The phase obtained from the raw CSI measurements is corrected by first unwinding it and then applying a linear transformation. The main idea is to remove the terms δt and β by considering the phase across the whole frequency band [27].

B. Noise Reduction

Since raw CSI data is noisy in nature, we adopt the DWT technique to filter out in-band noise and preserve the high frequency components, thereby introducing less distortion to the signal. DWT-based noise filtering consists of transforming the signal into the wavelet domain whereby the signal is divided into several frequency levels called wavelets that consist of the detail and approximation coefficients [29]. These can be mathematically represented as [5]:

$$y_{1,\text{low}}[n] = \downarrow Q\left[\sum_{k=-\infty}^{\infty} x[k]g[n-k]\right], \tag{5}$$

$$y_{1,\text{high}}[n] = \downarrow Q\left[\sum_{k=-\infty}^{\infty} x[k]h[n-k]\right], \tag{6}$$

where $y_{1,\text{low}}[n]$ and $y_{1,\text{high}}[n]$ are the approximation and detail coefficients, respectively, k denotes the frequency index, $x[k]$ is the input signal, $\downarrow Q[\cdot]$ represents a downsampling filter, $g[n]$ is a low-pass filter and $h[n]$ is a high-pass filter. The highest wavelet level is considered as noise. For each level, the noise and threshold for that level are estimated. The threshold is adapted for lower wavelets and the noise is removed in all levels without introducing significant distortion to the signal. In addition to DWT denoising, 1-D median filtering is also applied to the signal to remove any unwanted transients or spikes in the signal, especially when no activities were performed and the signal should be stable in this case.

C. Data Reduction

The raw CSI measurements were collected on a device equipped with the Intel 5300 NIC with three receiving antennas. For each pair of transmitting antenna and receiving antenna, we obtain CSI values from 30 OFDM subcarriers using the Linux CSI tool [28]. Therefore, if we have one transmit and three receive antennas, we obtain $1 \times 3 \times 30 = 90$ complex CSI values for each packet. The packet sampling rate was set at 1 kHz and hence in one second, we obtain 1000 packets each of size 90. This results in a large amount of data that needs to be processed and fed to a learning algorithm for classification. Therefore, dimension reduction is necessary for a SENS system. In this work, the popular PCA method has been used to identify the time-varying correlations between CSI streams which are then optimally combined to extract components that represent the variations caused by human activities. The number of principal components is empirically selected to achieve a good trade-off between classification performance and computational complexity [12]. Following DWT denoising, the first two or three principal components are sufficient to capture most of the variance in the CSI data stream [29]. Similar to [12], in the SENS system, we extract the first six principal components. However, the first one is safely discarded since it contains a lot of noise and will not result in any loss of information [2], [12], [30]. Therefore, only the next five principal components are retained for further processing.

D. Doppler Spectrogram Generation

CSI measurement is highly sensitive to the surrounding environment and RF reflections from the human body exhibit different frequencies when performing different activities. These frequencies can be distinguished in the time-frequency domain (spectrogram) by applying STFT to the PCA-denoised signal. Basically, the STFT applies a sliding window to obtain equally-sized segments of the signal and then performs FFT on the samples in each segment. The STFT of a time-domain input signal, $x[n]$, is given as:

$$X(t,k) = \sum_{n=-\infty}^{\infty} x[n]w[n-t]e^{-jkn}, \tag{7}$$

where t and k denote time and frequency indices, respectively, and $w[n]$ represents a window function (e.g., Hamming window). The spectrogram has three dimensions, namely time, frequency, and FFT amplitude. The Doppler spectrogram from STFT identifies the change of frequencies over time. The window size for FFT determines the trade-offs between frequency and time resolutions. For instance, a larger window size results in a higher frequency resolution but lower time resolution. The spectrograms are generated from the five principal components which are then averaged to obtain the final spectrogram. Unlike Doppler radar, the CSI spectrogram does not associate negative frequencies and hence the direction information is not available.

V. Signal Processing for PWR System

A. Cross Ambiguity Function

A PWR system consists of two synchronized receiver channels; a surveillance channel $x(t)$ which measures the target signals from the monitoring area, and a reference channel $y(t)$ which records the signal from the Wi-Fi access point. CAF processing has been used to obtain the range τ and Doppler f_d parameters by taking the FFT of the cross-correlated signals from the surveillance and reference channels. Doppler resolution is defined by the integration of time T_i as: $\Delta f_d = 1/T_i$. This allows the Doppler resolution to be adjusted for detecting human activities. The CAF equation can be written as:

$$CAF(\tau, f_d) = \int_0^{T_i} x(t)y^*(t-\tau)e^{j2\pi f_d t}dt, \tag{8}$$

where $*$ denotes a complex conjugate operation. Equation (8) requires a high computational load due to the long FFT which is not suitable for real-time processing in our system. Thus, the batch processing [19] has been used for complexity reduction. This is achieved by dividing a long sequence into several short batches so that the cross-correlation and FFT processes are faster. The CAF with batch processing can be expressed as:

$$CAF(\tau, f_d) = \sum_{n=0}^{N_b-1} \int_0^{T_b} x_n(t)y_n^*(t-\tau)e^{j2\pi f_d t}dt, \tag{9}$$

where N_b is the number of batches, T_b is the batch length and n is the index of the beacon. In order to obtain better performance, the reference channel was pointing towards the Wi-Fi AP in our experiments to make it free from interference due to human activities.

B. Direct Signal Cancellation

Note that the PWR system does not need to remove the SFO/STO as in the SENS system since both the surveillance and reference channels are synchronized through the USRP platform and hence they share the same clock source. A major drawback associated with PWR arises from the DSI component which undergoes perfect correlation with the reference signal, producing large range and Doppler sidelobes that can mask the weaker target echoes. Furthermore, the

TABLE II: System Implementation

System	SENS	PWR
Wi-Fi Signal	2.4 GHz (channel 1)	2.4 GHz (channel 6)
Hardware	Intel 5300 Wi-Fi [28]	NI USRP-2921 [31]
Subcarrier/bandwidth	30 (out of 56) subcarriers	1 MHz (out of 20 MHz)
Antenna	Omni-directional (6 dBi)	Directional (13 dBi)
Frame Rate	1000 per second	1000 per second
Measurement Rate	1000 Hz (same as frame rate)	10 Hz
Real-time Processing	No	Yes
Output Data Size per Second	90k: 1(tx) $\times$ 3(rx) $\times$ 30(sub carriers) $\times$ 1000(packets)	30k: 100(Doppler bin)$\times$30(range bin) $\times$ 10(sliding window)

DSI increases the dynamic range requirement of the system. However, angular nulling with the antenna and interference cancellation techniques in the receiver can be used to suppress the unwanted effects and improve system performance. A modified version of the CLEAN algorithm proposed in [17] is therefore adopted to suppress the DSI in our CAF processing. This CLEAN algorithm shares a similar structure to the CAF process but generates the self-ambiguity surface from the reference channel. This self-ambiguity surface is then used as an estimation of the direct signal, which is calculated as:

$$CAF^k(\hat{\tau}, \hat{f_d}) = CAF^k(\tau, f_d) - \alpha^k CAF_{self}(\tau - T_k, f_d), \tag{10}$$

where $CAF^k(\hat{\tau}, \hat{f_d})$ is the cleaned surface at the kth iteration, CAF_{self} is the self ambiguity surface, α^k and T_k are the amplitude and phase shift of maximum peak in the kth CAF surface. The CLEAN algorithm is implemented in the same way as the CAF process due to their similar structure.

C. Noise Reduction

After the CLEAN algorithm, we can still observe some noise in the CAF surface. One of the main reasons is that the CAF process over time gaps between Wi-Fi frames may introduce some noise. Furthermore, the CAF may be incorrectly processed due to strong interfering Wi-Fi signals from other APs or weak received signals from the desired AP. One common solution is to apply CFAR to estimate the background noise distributions as follows:

$$\Lambda = \frac{1}{N_\tau \cdot N_{f_d}} \sum_{i=1}^{R_\tau} \sum_{j=1}^{R_{f_d}} CAF(\tau_i, f_{d_j}), \tag{11}$$

where Λ is the threshold mapping for CAF. i and j are the indices for range and Doppler bins, respectively, N_τ and N_{f_d} are the training length in range and Doppler bins, respectively. This threshold mapping is then used for normalizing the power and removing the noise as $P(i,j) = |CAF(i,j)|^2/\Lambda$. $P(i,j) < 1$ implies no motion and the corresponding point in CAF is replaced with zeros. Otherwise, it is inferred that an activity has occurred.

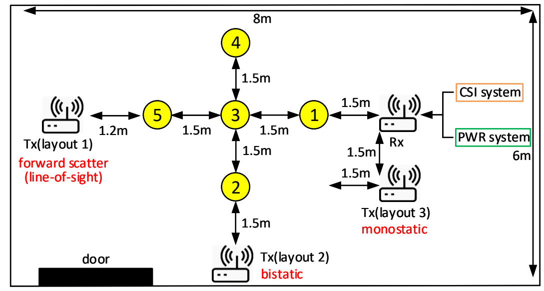

Fig. 4: Experiment layout

VI. System Implementation & Experiment

A. System Implementation

To enable a fair comparison between the SENS and PWR systems, we implemented both systems with almost the same system settings and Wi-Fi firmware configurations. Details of both systems are provided in Table II.

The Intel 5300 [28] NIC has been selected to implement the SENS system in this work. By default, CSI measurements from 30 out of 56 subcarriers can be extracted from each transmit-receive antenna pair and stored on the computing unit (a laptop) for off-line processing. The PWR system was built based on our previous work [11] with two synchronized NI USRP-2921 [31] as RF front-end by sharing one clock source. Then sampled data is transferred to the computing unit. It is not necessary to consider the entire bandwidth as we use the Doppler information. Therefore, a bandwidth of 1 MHz was found to be the trade-off between system performance and stability for our PWR system. Outputs from the two systems were timestamped from an external NTP time server for synchronization purposes.

Both systems were running in the 2.4 GHz Wi-Fi frequency band but on different channels to avoid cross-interference. This is because the SENS system continuously pinged the transmitter (Wi-Fi AP) to obtain the CSI packets. This two-way communication will introduce interference to the PWR system unless they operate on different channels. The frame rate was set at 1k per second in both systems to capture noticeable changes or patterns in the time domain signal which are caused by human motion. Measurement rate represents the number of system outputs per second. For the SENS system, measurement rate is defined as the number of received packets per second, which is 1 kHz. However, for the PWR system, the measurement rate is limited by the amount of baseband signals that can be processed by the computing unit. Based on empirical experience, the measurement rate of the PWR system was set at 10 Hz.

B. Experiment Layout

All measurements were carried out within an office area and the experiment layouts are illustrated in Fig 4. The monitoring area was approximately 8m x 6m with computers and office furniture in the surroundings. To compare the detection performance of the two systems with different geometries, the location of the receive antenna remained the same throughout,

TABLE III: Activity Description

	Activity	Description
(1)	walking	walking in the direction of 1-3-5, 2-3-4; this represents a long, high-level body movement
(2)	sitting	sitting to a chair at position, 1,2,3,4,5; this represents a short, medium-level body movement
(3)	standing	standing from a chair at position, 1,2,3,4,5; this represents a short medium-level body movement
(4)	laying	laying down to floor at position, 1,2,3,4,5; this represents a long low-level body movement
(5)	standing from floor	standing from floor at position, 1,2,3,4,5; this represents a long, low-level body movement
(6)	picking	picking up small items at position, 1,2,3,4,5; this represents a short, medium-level body movement

whereas the Wi-Fi transmitter (AP) was moved in each layout as per Fig. 4. Layout 1 refers to the scenario whereby the transmitter-object-receiver alignment is around 180 degrees. This forms a forward scatter geometry which is also known as the line-of-sight (LoS). Layout 2 is when the transmitter-object-receiver is around 90 degrees and this forms a bistatic geometry. Layout 3 is when the transmitter-object-receiver is less than 45 degrees and this is known as a monostatic geometry. Five testing positions were used during the experiments and they were separated by 1.5m from each other. These positions are used to evaluate the effect of the system geometry on the activity classification accuracy.

C. Dataset

In this pilot study, we conducted six day-to-day activities; namely, walking, standing from a chair, sitting on a chair, laying down on the ground, standing from the ground and picking up a small object from the ground. The walking activity covered 3 positions while this activity was performed whereas the other activities were performed at one specific position before moving to the next. Furthermore, we considered different transmitter-receiver layouts to cover both LoS and non LoS (NLoS) conditions, as would be the case in a real-world environment. The descriptions of the above activities are given in Table III. We applied a sliding window to the Doppler spectrograms and extract 4 seconds of Doppler data for each measurement, no matter the difference in the activity duration. Five volunteers (four males and one female) of different age groups (ranging from 22 to 30) were involved in the experiments. In this work, we have collected a total of 1,122 data samples from the six activities. Among these, layout 1 has 138 samples, layout 2 has 826 samples and layout 3 has 158 samples.

VII. Experimental Results

In this section, the activity recognition performance of both the SENS and PWR systems is presented. A simple 2D convolutional neural network (CNN) has been used as the classifier. The CNN includes one convolutional layer, one max-pooling layer and two fully connected layers. Since the input data size is different for the two systems, some parameters are different in the two classifiers.

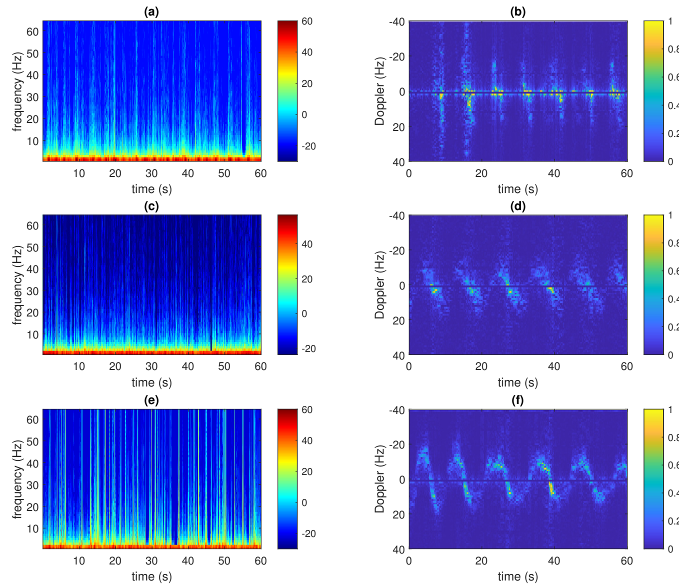

Fig. 5: Walking spectrogram obtained from (a) SENS system in layout 1, (b) PWR system in layout 1, (c) SENS system in layout 2, (d) PWR system in layout 2, (e) SENS system in layout 3 and (f) PWR system in layout 3

A. Spectrogram Comparison

Fig. 5 presents the difference in the walking spectrograms obtained from the SENS and PWR systems for all three layouts. Test subjects walked along with positions 2-5-8 repeatedly at a constant speed. As can be seen from Fig. 5, the subsequent spectrograms have similar signatures in the SENS system with a dominant high Doppler frequency which we attribute to movement of the torso, and small frequencies which are related to the movement of the limbs.

In comparison, Doppler signatures for walking in the PWR system present a significantly different footprint in terms of Doppler profile, shift and amplitude. This can be explained by the fact that the PWR system is highly sensitive to the geometry of the transmitter and receiver locations. The PWR spectrogram in layout 1 (Fig. 5(b)) shows a very low Doppler shift since the relative velocity between the transmitter-object and object-receiver is almost zero when the PWR system operates in LoS. The spectrogram in layout 2 (Fig. 5(d)) and layout 3 (Fig. 5(f)) have clearer Doppler signatures and more significant Doppler shifts.

In addition, spectrograms from SENS system do not contain information regarding the walking direction, whereas the sinusoidal wave in the PWR system clearly indicates its velocity and direction. This is because in the SENS system we only use CSI magnitude measurements. The direction information could be inferred by looking at the phase changes within the physical layer protocol data unit (PPDU) or across PPDUs. These details can be captured if one uses specialized equipment such as USRP. However, in inexpensive WiFi NICs, the extracted raw CSI measurements are affected by phase offsets as the hardware is far from ideal. Therefore, it is extremely challenging to correct the phase information in the

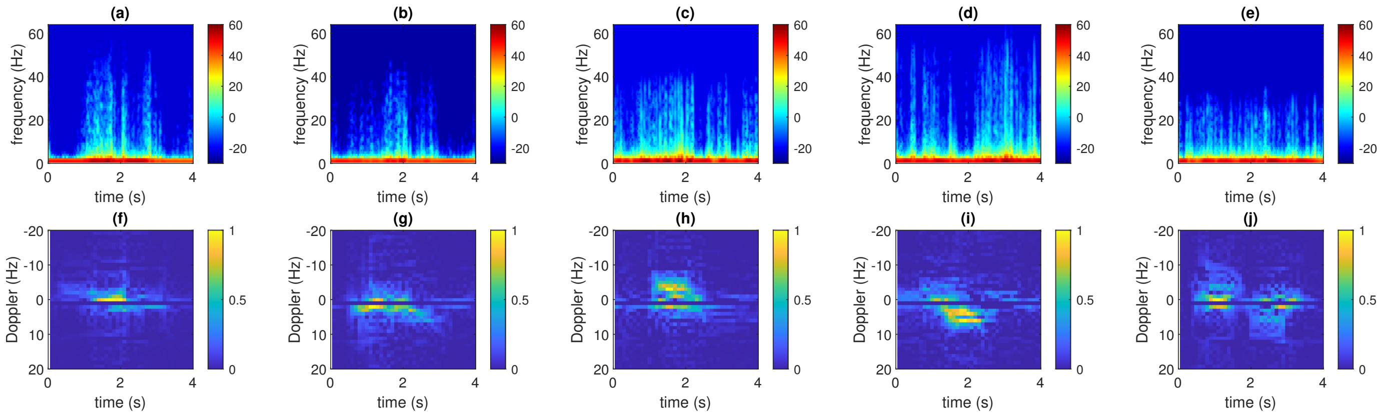

Fig. 6: Spectrogram obtained from layout 2 by SENS system: (a) sitting, (b) standing, (c) laying, (d) standing from floor, (e) picking and from PWR system: (f) sitting, (g) standing, (h) laying, (i) standing from floor, (j) picking

raw CSI data of COTS WiFi devices.

The PWR system has an integration time of 1 second which is sufficient to observe the direction of the object. However, the PWR system is less sensitive to micro Doppler when capturing a large movement, for example, the limbs' Doppler during walking. This is because the dominant Doppler pulse can easily mask the micro Doppler pulses.

Fig. 6 presents spectrograms for the other five activities as processed by the two systems. As it can be observed in Fig. 5, the frequency shifts in the CSI's spectrograms are lower than those in the walking spectrogram. Generally, all frequency shifts or Doppler shifts in Fig. 6 are lower than those in Fig. 5 due to the relatively slower body motion. There are some lower frequency shifts in the CSI spectrograms which relate to part body movement. For example, we observe short and weak frequency shifts from the "standing from chair" activity spectrogram (Fig. 6(a)), while the "picking up" activity (Fig. 6(e)) has the lowest frequency shift.

There are more patterns that can be observed in the PWR's spectrograms for certain activities. For example, "sitting on a chair" (Fig. 6(f)) and "laying down on the floor" (Fig. 6(h)) spectrograms both have a negative Doppler shape, since both activities consist of a downward body movement. This trend can also be observed from the "standing from chair" (Fig 6(g)) and "standing from floor" (Fig 6(i)) spectrograms, where both contain a positive Doppler shape. The "picking up" activity contains two-part movements, bending over and straightening up the body. As expected, we can see a negative Doppler shape followed by a positive shape (Fig. 6(j)).

B. Classification Accuracy Versus Activity

We first conduct the classification results for all activities in terms of different positions or layouts. 80% of the dataset was chosen randomly and used for training, and the remaining 20% was used for testing. The overall accuracy for the SENS system is 67.3% and the PWR system has almost similar accuracy at 66.7%. These accuracies are lower than those achieved in studies like [2], [11], [15], [32] (more than 90% in accuracy). The reason for the low accuracy is because of the mixture of forward scatter (LoS), bistatic and monostatic

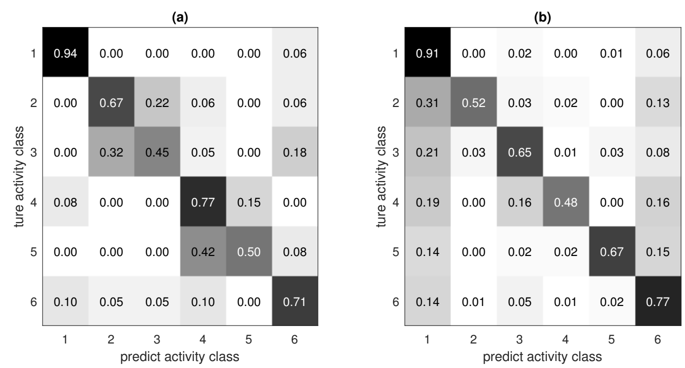

Fig. 7: Classification matrices for activity recognition for combined layouts 1,2,3 in (a) SENS system and (b) PWR system

(NLoS) layouts that result in different Doppler signatures as presented in Fig. 5. Also, the change of measurement position means the variation in signal reflection power at the receiver side would cause the strength of the Doppler signal to become unstable. Nevertheless, this accuracy is still acceptable and can be considered to be a benchmark when different physical layouts and positions are mixed up together.

The confusion matrices for the SENS system and PWR system are shown in Fig. 7(a) and Fig. 7(b), respectively. As can be seen, both systems have the best classification result for activity 1 (walking), where the accuracy is more than 90%. This is because the walking activity contains higher Doppler shifts than other activities in any directions or layouts. The second best result is observed for activity 6 (picking up), with accuracy over 70%. The other four activities have relatively low accuracy. The SENS system has the worst performance for activity 3 (standing) and activity 5 (standing from floor), whereas the PWR system has the worst performance for activity 2 (sitting) and activity 4 (laying down). Moreover, the wrong predictions in the SENS system mostly happen between the pair of activities like "sitting on a chair" and "standing from a chair", "laying down" and "standing from floor". The reason is that the SENS system measures Doppler

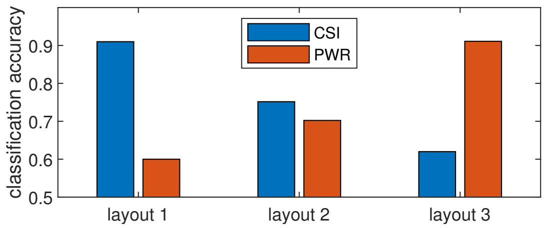

Fig. 8: Classification versus three layouts

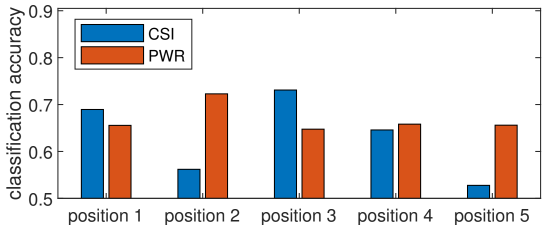

Fig. 9: Classification versus different positions

shift in a short time and is therefore more sensitive to activities with different time duration. In comparison, most incorrect predictions in the PWR system occur for the walking activity. This is because the PWR system has a longer integration time (1s in this work), so that long duration activities are easier to distinguish. This accuracy could easily be improved by choosing the appropriate layout for each system.

One of the major factors that affect the recognition accuracy is the geometry of transmission and reception. Thus, it is interesting to evaluate the activity recognition accuracy in different physical setups. To evaluate such performance, both the training and testing data were used within the same layout. The results are shown in Fig. 8. As expected, the SENS system has the best performance in layout 1 at 91% and worst performance in layout 3 at 62%, whereas the PWR system has the best performance in layout 3 with an accuracy of 91.1% and worst in layout 1 with an accuracy of 60%. Both systems have almost similar accuracy in layout 2 around 70%, which is more than the accuracy in Fig. 7. As mentioned previously, the SENS and PWR systems have different mechanisms in processing the Wi-Fi signal. The SENS system has better performance in the forward scatter (LoS) layout while the PWR system has better performance in the monostatic layout. These results demonstrate the coverage sensitivities of the two systems that can be used in real applications.

Next, we calculate the accuracy over each position shown in Fig. 4. In this experiment, we tested the data for a specific position and trained the data for all other positions (excluding the walking activity which covers several positions). The classification accuracy for each position is shown in Fig. 9. As can be observed, the two systems differ in their position accuracy. More specifically, the SENS system has the worst performance at positions 2 and 5, where the accuracy is below 60%. The SENS system works best at position 3 which is close to a LoS layout. In comparison, the PWR system has a more balanced performance across all positions since the bistatic angle is relatively similar. The spectrograms in positions 2 and 5 have relatively similar Doppler signatures. These results

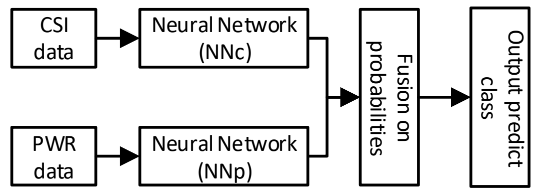

Fig. 10: Fusion framework for two systems

TABLE IV: Combined Accuracy

Dataset	Method	SENS	PWR	Combined
layout 2	Addition	75.7%	72.8%	79.8%
layout 2	Multiplication	75.7%	72.8%	74.0%
layout 1,2,3	Addition	67.3%	66.7%	74.2%
layout 1,2,3	Multiplication	67.3%	66.7%	70.2%

show that SENS and PWR systems have slightly different coverage, and therefore a fusion of the two systems could improve the performance in weaker positions.

C. Combined Classification accuracy

So far, classification results are calculated separately for the two systems. As discussed before, there are considerable differences in classification accuracy in terms of system layout and position. Here, we combine the results from both systems to further improve the accuracy of Wi-Fi sensing. Inspired by the work in [33], a simple fusion framework has been used, as shown in Fig 10. Here we export the probabilities, P_c and P_p, for each activity from the Neural Networks used for the SENS and PWR systems, respectively. We set the two systems with same weight, and use two methods to calculate the combined probabilities, P_f; the addition method where $P_f = \frac{1}{2}(P_c+P_p)$ and the multiplication method where $P_f = P_c * P_p$.

The combined accuracy for layout 2 and layouts 1,2,3 are given in Table IV. We followed the same procedures as discussed in Section VII-B. As can be seen, there are some improvements in the combined accuracy as compared to the accuracy obtained from the SENS and PWR systems separately. In layout 2, the combined accuracy using the addition method is 79.8% which is 4.1% and 6.0% higher than the accuracy of each individual system, whereas the multiplication method results in a slightly lower combined accuracy. The combined accuracy shows even better improvement in layout 1,2,3 which achieves 74.2%. In addition, these improvements are generated using a simple fusion framework based on the probabilities from the two systems. It is envisioned that a more robust fusion process using the CSI and PWR data could further improve the classification accuracy.

For layout 1 and layout 3, the imbalanced performance between the two systems makes the fusion process ineffective. In some cases, the fusion process could lower the original accuracy since we consider equal weights in this work.

VIII. Discussion

This section discusses the challenges that we faced during the implementation and experimentation with the two systems.

Future improvements to the hardware configurations for an integrated measurement system, signal processing and machine learning algorithms are also discussed.

A. Resilience to Environmental Changes

Wi-Fi signals are very sensitive to various factors such as the geometry of transmission and reception, environmental conditions and operational parameters of the communication network. It is crucial to build a robust Wi-Fi sensing system that can be adapted for different environments and Wi-Fi AP settings (e.g., bandwidth, transmit power, MIMO capability, etc.) but this represents a challenge. For example, the direction and orientation of the person with respect to the Wi-Fi AP and receiver can change continuously. The distance between the person and Wi-Fi AP could also be varied. In practical scenarios, there may be multiple people or other moving objects around that could block the reference channel as well as the baseline (LoS) between the transmitter and receiver. It is very challenging for Wi-Fi sensing systems to have the generalization ability to automatically adapt to new and unseen data. In other words, a Wi-Fi sensing system should also work when the device is placed in a new environment, unknown location and operate for new talents.

For a SENS system, it requires a process to understand the surrounding environment during a static measurement which may be hard to operate in real scenarios. One solution is to convert the CSI measurement into Doppler spectrograms to calculate the change in frequency. However, this does not entirely solve the problem as it is extremely challenging to correct the phase in the raw CSI data due to the imperfection in hardware and non-synchronized sampling clocks and frequencies between the transmitter and receiver. In comparison, the PWR system directly outputs a Doppler spectrogram which is less sensitive to the static objects and previously trained model can be used in a new environment. However, the PWR system needs to overcome the challenge where two channels are required. The re-creation of the transmitted signal should be improved through a robust algorithm instead of using a reference channel.

B. Efficiency in Spectrum Usage

The fundamental purpose of Wi-Fi is for wireless communications. Sensing is a peripheral application that can either be used to optimize the performance and quality of service (QoS) of the network, or secondary applications in healthcare, IoT, security etc. The majority of previous studies which have investigated CSI-based sensing systems [8], [12], [14], [34], use a high frame rate to achieve good performance. However, high frame rate can be regarded as the exchange of redundant information which occupies a considerable amount of the already-limited Wi-Fi spectrum. This in turn affects the network performance, degrading the QoS for connected users. Moreover, sending unnecessary packets for CSI measurements influences not only the measuring device but also the nearby Wi-Fi devices, since the packets occupy Wi-Fi resources in both time and frequency domains. In contrast, the passive nature of the PWR system means that no extra packets are transmitted for sensing purposes. This minimizes the influence on communication systems, but the PWR's performance is highly dependent on the density of the Wi-Fi packages which might be a problem when the data traffic through the AP is low.

In addition, SENS system does not take full advantage of a Wi-Fi packet. Recall from Fig. 2, the SENS system only uses the preamble signal to obtain the desired CSI but does not have information about the transmitted data signal. Despite that the PWR system can capture the whole packets, however, it also captures the time gaps between packets which are redundant for sensing and this results in a high computational processing overhead. To enhance the detection performance, it is important to maximize the usage of the Wi-Fi package while filtering out the time gap period. This is required for the data signal generation in the PWR system using the reference channel method.

C. Beamforming

The latest IEEE 802.11ac standard use the beamforming technique which could have an adverse impact on both SENS and PWR sensing as it changes the amplitude and phase of the Wi-Fi signals. As a result, the CSI measurements may become unstable and difficult to process if the beamforming matrix is not available at the receiver side. The PWR system faces more challenges due to the beamforming technique. Traditionally, passive radar operates with relatively low bandwidth and uses a single carrier signal like FM radio and analog television. Multiple antennas in the beamforming technique means the aquisition of the PWR's reference channel becomes even more complicated. Acquiring the reference channel using a single directional antenna from a MIMO AP will be challenging since each received signal will have a different amplitude and phase. The variation in phase difference may generate erroneous Doppler pulse in the CAF surface and cause similar sidelobe problem in the PWR system. Nonetheless, beamforming can be advantageous for Wi-Fi sensing by providing spatial information in addition to the Doppler and range information. However, current SENS and PWR systems have not used this new technique to generate joint spatial and Doppler data.

D. Challenges in Signal Processing

Using commercial NICs, SENS systems can obtain fine-grained CSI measurements directly without further processing. However, the size of the CSI measurements (shown in Table II) is proportional to the number of antennas and frame rate. This means a significant computational power is required to process such amount of data, although it is possible to reduce the size of the data using techniques such as PCA, which captures most of the variance among the subcarriers over multiple antennas in only a few principal components. On the other hand, the raw CSI measurement is too noisy to be used directly for sensing purposes and hence the SENS signal processing represents a very important engineering task. The processing of CSI measurements to obtain meaningful information such as Doppler, range, AoA, ToF, etc, is necessary and it is worthwhile to develop algorithms that are useful for joint activity recognition and localization applications.

From the Doppler spectrograms, we realize that the traditional CAF process (Equation 8) in the PWR system could not deliver sufficient range resolution for human sensing due to the limited Wi-Fi bandwidth. Also, the integration time (one second in this work) which defines the Doppler resolution, is too long for activities consisting of hand gestures. It is believed that a more efficient CAF processing with time synchronization (to extract effective Wi-Fi signal) could further improve the PWR system in both range and Doppler resolutions. Moreover, the SENS system has a low sensitivity to activities performed far from the baseline while the PWR system has a low sensitivity to activities performed close to the baseline. Thus, information fusion from both systems could significantly improve the coverage for Wi-Fi sensing.

E. Challenges in Machine Learning Algorithms

Machine learning algorithms in Wi-Fi sensing face several challenges. Firstly, the training data available for some activities such as falling down (especially in elderly people) are difficult to collect and may be insufficient to train a model due to under-fitting. This is a class imbalance problem [35], where most standard classifier learning algorithms assume a relatively balanced class distribution. Such a situation represents a challenge in current Wi-Fi sensing works [11], [12], [15] and thus a different approach [35] is required for the imbalanced activity classes.

Secondly, a large dataset is required to properly train a classifier, taking into account various factors like transmit/receive geometry, abnormal activities and different height/weight of people which could potentially change the Doppler pattern for a given activity. This may not be feasible since the data collection process will be time consuming and may incur a high cost. However, two common solutions are available, namely, model-based algorithms such as Finite Difference Time Domain (FDTD) [36] which studies the physical theories or statistical model of the target, and learning-based algorithms such as generative adversarial network (GAN) [37] which generates new datasets based on a pre-trained network. Some early works like [37], [38] have shown the potential of using generated Doppler spectrum to improve the accuracy in activity recognition. However, current works applying these algorithms are still in the early stage and they focus on simple activities performed mostly in a static environment (controlled experiments).

Another challenge is the cross-device/sensor in Wi-Fi sensing. Multiple Wi-Fi devices can be combined together to achieve higher performance and efficiency. Due to the rapidly increasing demand in wireless data, there will be more Wi-Fi devices available in different scenarios. These devices are location separated which could provide extra information for cross-device sensing. In addition to Wi-Fi devices, many other types of sensors such as cameras, mobile phones, laptops, IoT devices, etc., can be used for cross-sensor sensing. The latter can reduce human efforts for training machine learning algorithms. For example, video cameras can be used to generate automatic ground truth labels for the SENS and PWR systems.

IX. Conclusions

In this paper, we presented and compared methods based on CSI (SENS) and radar (PWR) protocols for activity sensing using Wi-Fi transmissions. We investigated the difference between these two systems in terms of fundamental working principles and key challenges. We report on a range of human activity data obtained from these two systems in a realistic indoor environment and compare the classification accuracy in terms of system and surveillance area geometries. The SENS and PWR systems show the best performance in the line-of-sight and monostatic layouts, respectively. Moreover, we have demonstrated that a fusion process on both systems could easily improve the accuracy of activity recognition. Future work includes the development of a more robust system that can combine the advantages of the SENS and PWR systems. Also, the efficiency in spectrum usage and beamforming technique is worth considering in Wi-Fi sensing.

Acknowledgments

This work was funded under the OPERA Project, the UK Engineering and Physical Sciences Research Council (EPSRC), Grant EP/R018677/1.

References

[1] C. J. Caspersen, K. E. Powell, and G. M. Christenson, "Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research." *Public health reports*, vol. 100, no. 2, p. 126, 1985.

[2] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 363–373.

[3] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2016.

[4] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous wifi-based gesture recognition system," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2015, pp. 1472–1480.

[5] Y. Ma, G. Zhou, and S. Wang, "Wifi sensing with channel state information: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, p. 46, 2019.

[6] C. Feng, W. S. A. Au, S. Valaee, and Z. Tan, "Received-signal-strength-based indoor positioning using compressive sensing," *IEEE Transactions on mobile computing*, vol. 11, no. 12, pp. 1983–1993, 2011.

[7] Y. Shu, Y. Huang, J. Zhang, P. Coué, P. Cheng, J. Chen, and K. G. Shin, "Gradient-based fingerprinting for indoor localization and tracking," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2424–2433, 2015.

[8] C. Yang and H.-R. Shao, "Wifi-based indoor positioning," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 150–157, 2015.

[9] A. Makki, A. Siddig, M. Saad, and C. Bleakley, "Survey of wifi positioning using time-based techniques," *Computer Networks*, vol. 88, pp. 218–233, 2015.

[10] W. Li, B. Tan, and R. J. Piechocki, "Non-contact breathing detection using passive radar," in *2016 IEEE International Conference on Communications (ICC)*. IEEE, 2016, pp. 1–6.

[11] W. Li, B. Tan, Y. Xu, and R. J. Piechocki, "Log-likelihood clustering-enabled passive rf sensing for residential activity recognition," *IEEE Sensors Journal*, vol. 18, no. 13, pp. 5413–5421, 2018.

[12] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*. ACM, 2015, pp. 65–76.

[13] K. Joshi, D. Bharadia, M. Kotaru, and S. Katti, "Wideo: Fine-grained device-free motion tracing using {RF} backscatter," in *12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15)*, 2015, pp. 189–204.

[14] S. Tan and J. Yang, "Wifinger: leveraging commodity wifi for fine-grained finger gesture recognition," in *Proceedings of the 17th ACM international symposium on mobile ad hoc networking and computing*. ACM, 2016, pp. 201–210.

[15] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial wifi devices," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1118–1131, 2017.

[16] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, T. Gu, and B. Xie, "Human respiration detection with commodity wifi devices: do user location and body orientation matter?" in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 25–36.

[17] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1218–1226, 2011.

[18] F. Colone, P. Falcone, C. Bongioanni, and P. Lombardo, "Wifi-based passive bistatic radar: Data processing schemes and experimental results," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1061–1079, 2012.

[19] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive wifi doppler-radar and its applications," in *2014 International Radar Conference*. IEEE, 2014, pp. 1–6.

[20] S. Di Domenico, G. Pecoraro, E. Cianca, and M. De Sanctis, "Trained-once device-free crowd counting and occupancy estimation using wifi: A doppler spectrum based approach," in *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2016, pp. 1–8.

[21] W. Li, B. Tan, and R. Piechocki, "Opportunistic doppler-only indoor localization via passive radar," in *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. IEEE, 2018, pp. 467–473.

[22] Intel wi-fi link 5300. [Online]. Available: https://www.intel.co.uk/content/www/uk/en/products/docs/wireless-products/ultimate-n-wifi-link-5300-brief.html

[23] N. Tadayon, M. T. Rahman, S. Han, S. Valaee, and W. Yu, "Decimeter ranging with channel state information," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3453–3468, 2019.

[24] B. Tan, K. Woodbridge, and K. Chetty, "A wireless passive radar system for real-time through-wall movement detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2596–2603, 2016.

[25] M. Cherniakov and D. Nezlin, *Bistatic radar: principles and practice*. Wiley Online Library, 2007.

[26] N. J. Willis, *Bistatic radar*. SciTech Publishing, 2005, vol. 2.

[27] X. Dang, X. Tang, Z. Hao, and Y. Liu, "A device-free indoor localization method using csi with wi-fi signals," *Sensors*, vol. 19, no. 14, p. 3233, 2019.

[28] "Linux 802.11n CSI tool," https://dhalperi.github.io/linux-80211n-csitool/, (Accessed on 05/11/2020).

[29] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "Falldefi: Ubiquitous fall detection using commodity wi-fi devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–25, 2018.

[30] S. Liu, Y. Zhao, F. Xue, B. Chen, and X. Chen, "Deepcount: Crowd counting with wifi via deep learning," *arXiv preprint arXiv:1903.05316*, 2019.

[31] "Ni usrp 2921," http://sine.ni.com/nips/cds/view/p/lang/en/nid/212995, (Accessed on 05/11/2020).

[32] S. Duan, T. Yu, and J. He, "Widriver: Driver activity recognition system based on wifi csi," *International Journal of Wireless Information Networks*, vol. 25, no. 2, pp. 146–156, 2018.

[33] P. Zappi, T. Stiefmeier, E. Farella, D. Roggen, L. Benini, and G. Troster, "Activity recognition from on-body sensors by classifier fusion: sensor scalability and robustness," in *2007 3rd international conference on intelligent sensors, sensor networks and information*. IEEE, 2007, pp. 281–286.

[34] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring motion direction using commodity wi-fi for interactive exergames," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 1961–1972.

[35] Y. Sun, A. K. Wong, and M. S. Kamel, "Classification of imbalanced data: A review," *International journal of pattern recognition and artificial intelligence*, vol. 23, no. 04, pp. 687–719, 2009.

[36] J. B. Schneider, "Understanding the finite-difference time-domain method," *School of electrical engineering and computer science Washington State University*, p. 181, 2010.

[37] K.-S. Zheng, J.-Z. Li, G. Wei, and J.-D. Xu, "Analysis of doppler effect of moving conducting surfaces with lorentz-fdtd method," *Journal of Electromagnetic Waves and Applications*, vol. 27, no. 2, pp. 149–159, 2013.

[38] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Gan-based synthetic radar micro-doppler augmentations for improved human activity recognition," in *2019 IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–5.

Dr. Wenda Li received the M.Eng. and Ph.D. degree in Electrical and Electronic Engineering from the University of Bristol in 2013 and 2017 respectively. He worked at University of Birmingham as a Research Fellow before joining University College London. He is currently working as a Research Fellow in the Department of Security & Crime Science at University College London. His research focuses on the signal processing for passive radar and high-speed digital system design for wireless sensing applications in healthcare, security and positioning. His research in passive WiFi radar has led to a number of IEEE conference and journal publications.

Mohammud Junaid Bocus received the B.Eng. degree (first-class honors) in Electronic and Communication Engineering from the University of Mauritius, Mauritius in 2012, the M.Sc. (Distinction) degree in Wireless Communications and Signal Processing from the University of Bristol, Bristol, U.K, in 2015 and PhD degree in Electrical and Electronic Engineering from the University of Bristol, Bristol, U.K, in 2020. His research interests include terrestrial and underwater wireless communications, video coding, computer vision and machine/deep learning. He is currently working as a research associate at the University of Bristol, focusing on concurrent passive activity recognition and localization using commercial off-the-shelf Wi-Fi and ultra-wideband (UWB) systems.

Chong Tang is now a PhD student in the UCL Department of Security and Crime Science (SCS), where he is invetigating to apply Passive WiFi Radar (PWR) system for occupancy detection and reconstruct human-skeletal model from Doppler-spectrogram-only data. He has received a Bachelors degree in Automation from Sichuan University and Electrical & Electronic Engineering from University of Nottingham in 2018, and a Master degree in Robotics from UCL in 2019.

Robert Piechocki is a full Professor in the School of Computer Science, Electrical and Electronic Engineering and Engineering Maths, University of Bristol. He is also a Fellow at The Alan Turing Institute. His research interests span the areas of Connected Intelligent Systems, Wireless Networks, Information and Communication Theory, Statistics and Machine Learning. His domain expertise is Connected and Automated Mobility (CAM) and wireless sensing for eHealth. In his research work he strives to develop solutions for decision making and inference in networked systems which communicate over resource constrained and unreliable links. Rob has published over 200 papers in peer-reviewed international journals and conferences and holds 13 patents in these areas.

Karl Woodbridge is Emeritus Professor of Electronic and Electrical Engineering at University College London. Recent research interests include multi-static and software-defined radar systems, passive wireless surveillance and Doppler classification using machine learning methods. Current research is focussed on the development of passive wireless based sensors for activity detection and classification with application to Healthcare, IoT and Security. He is a Fellow of the IET, a Fellow of the UK Institute of Physics and a Senior Member of the IEEE. He has served on technical and organising committees for a wide range of International conferences and published or presented over 250 journal and conference papers in the areas of semiconductors, photonics and RF sensor systems.

Dr. Kevin Chetty is an Associate Professor at University College London (UCL) where he leads the Urban Wireless Sensing Lab. He has pioneered work in passive WiFi sensing; an area of radar research which is expected to facilitate advances in ubiquitous sensing and smart environments. His work in this area also covers new waveform designs for integrated communications and sensing and developing reference-free sensing approaches for passive radar systems. Additionally, he has research interests in using radar technology for behaviour classification through the exploitation of micro-Doppler signatures and machine learning techniques. Dr. Chetty is an author of over 70 peer reviewed publications and has been an investigator on grants funded by both government and industry.