Training-Free Human Vitality Monitoring Using Commodity Wi-Fi Devices

XIANG LI and DAQING ZHANG*, Peking University, China
JIE XIONG, University of Massachusetts, Amherst, USA
YUE ZHANG, SHENGJIE LI, YASHA WANG, and HONG MEI, Peking University, China

Device-free sensing using ubiquitous Wi-Fi signals has recently attracted lots of attention. Among the sensed information, two important basic contexts are (i) whether a target is still or not and (ii) where the target is located. Continuous monitoring of these contexts provides us with rich datasets to obtain important high-level semantics of the target such as living habits, physical conditions and emotions. However, even to obtain these two basic contexts, offline training and calibration are needed in traditional methods, limiting the real-life adoption of the proposed sensing systems. In this paper, using the commodity Wi-Fi infrastructure, we propose a training-free human vitality sensing platform, WiVit. It could capture these two contexts together with the target's movements speed information in real-time without any human effort in offline training or calibration. Based on our extensive experiments in three typical indoor environments, the precision of activity detection is higher than 98% and the area detection accuracy is close to 100%. Moreover, we implement a short-term activity recognition system on our platform to recognize 4 types of actions, and we can reach an average accuracy of 94.2%. We also take a feasibility study of monitoring long-term activities of daily living to show our platform's potential applications in practice.

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

Additional Key Words and Phrases: Wi-Fi, Device-free, Training-free, Human vitality

ACM Reference Format:
Xiang Li, Daqing Zhang, Jie Xiong, Yue Zhang, Shengjie Li, Yasha Wang, and Hong Mei. 2018. Training-Free Human Vitality Monitoring Using Commodity Wi-Fi Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 121 (September 2018), 25 pages. https://doi.org/10.1145/3264931

1 INTRODUCTION

In recent years, Wi-Fi based device-free sensing, which does not require attaching any device to the target, has attracted a lot of attention from researchers. Compared with other device-free sensing technologies such as camera [7, 26] and ultrasound [24, 41], Wi-Fi devices are ubiquitous in indoor environment so that we do not need to deploy any extra infrastructure. As reported in [44], there are already more than eight devices on average with Wi-Fi chipset embedded in a typical US home environment. Moreover, camera-based solutions raise severe privacy concerns in indoor environments while ultrasound-based methods have a very small coverage

*This is the corresponding author

Authors' addresses: Xiang Li, Daqing Zhang, Yue Zhang, Shengjie Li, and Hong Mei, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China; email: {lixiang13, dqzsei, zy.zhangyue, lishengjie, meih}@pku.edu.cn. Jie Xiong, University of Massachusetts, Amherst, Massachusetts, USA; email: jxiong@cs.umass.edu. Yasha Wang, Key Laboratory of High Confidence Software Technologies (Ministry of Education), National Engineering Research Center for Software Engineering, Peking University, Beijing, China; email: wangyasha@pku.edu.cn.

2474-9567/2018/9-ART121 $15.00
https://doi.org/10.1145/3264931

area and the performance degrades significantly in noisy environments. These disadvantages limit their real-life applications. Therefore, Wi-Fi has been considered particularly promising for ubiquitous indoor human sensing.

In the last few years, a lot of Wi-Fi based device-free human sensing applications have been proposed, such as indoor localization [20, 38, 50], activity recognition [39, 42], intrusion detection [19] and respiration monitoring [22, 35, 54], etc. For most of these applications, capturing the human vitality information (i.e. the human is *still* or *non-still* and in which area the human is staying) is essential. For example, respiration sensing with Wi-Fi needs to know when the target is still and then to monitor the respiration because body movements severely interfere with the fine-grained respiration sensing. On the contrary, activity recognition needs to know when the human is non-still and then to segment and recognize different activities. Moreover, a specific type of human activity usually takes place in a particular area, such as sleeping activity takes place in the bedroom and eating activity usually happens at the dining table. Thus, for activity recognition, knowing the rough area information can significantly help reduce the training size and increase the sensing accuracy. Furthermore, the long-term vitality information can be employed to infer a lot of useful high-level semantics about the target. For example, frequent toilet visits and a very short sleep during the night different from the target's usual routine are signs for medical attentions.

In this paper, we use *vitality* to represent the information including in which area the target is staying and whether the target is still or non-still. Although some approaches have been proposed to detect when the human is non-still or locate the human target in the indoor environment based on the Wi-Fi signal, there are several limitations preventing us from applying these approaches to obtain the human vitality information in practice. Existing systems usually detect whether the human target is non-still based on the variance or correlation of Wi-Fi signal in time domain [19, 28, 39, 42, 47]. These methods require significant amount of human efforts in offline training and calibration to learn the difference between static environment (i.e. the target is still) and dynamic environment (i.e. the target is non-still). To locate the target, most of existing systems employ fingerprint-based solutions [1, 38, 48], which require labor-intensive offline training to build the fingerprint database for localization. Angle-of-Arrival (AoA) based device-free localization solutions [20, 21] require careful phase calibration to remove the random phase offset between two RF ports during startup. To the best of our knowledge, there is still no such a platform which could detect when the target is non-still and in which area the target is staying using cheap commodity Wi-Fi devices without any human effort in offline-training or calibration.

In this paper, we design WiVit, a device-free human vitality monitoring platform based on commodity Wi-Fi devices, which could capture when the target is non-still and in which area the target is staying without human intervention. In a typical home environment, there are one Wi-Fi access point and multiple Wi-Fi-enabled home appliances such as TV, air conditioner, refrigerator, etc. WiVit utilizes the access point and the Wi-Fi-enabled devices to form transmission pairs. WiVit is composed of two key components. First, the activity detection module could detect human activities without offline-training or calibration. The key insight is that in an indoor environment, Wi-Fi signals not only propagate along the direct path to the receiver, but also are reflected by objects such as walls, furniture and the human body. The signal at the receiver is the superposition of signals from all paths. This phenomenon is called multipath propagation. When the human target is not still, no matter walking or just waving hands in-place, the target's movements will cause changes in the length of the target's reflected path. During a short-time period, the path length change speed can be considered as a constant. The phase readings of the reflected path signal also change at a constant speed accordingly. However, other path signals reflected by the furniture and walls do not change. Thus, by capturing this constant phase change and estimating the corresponding human reflected path change speed, which is only caused by the human movement and not affected by other multipath signals, the human activity can still be clearly detected even when complex multipath propagations exist. Second, the area detection module could identify which area the human is staying in without any human intervention. The basic idea of this module is that we employ multiple Wi-Fi transceiver pairs to divide the sensing space into multiple areas according to the geographical layout of the home environment. Each pair of

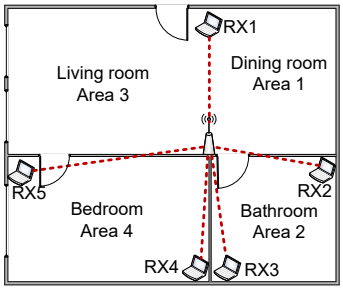


Fig. 1. An example of WiVit platform deployment.

transceivers is the boundary of two adjacent areas. Figure 1 shows an example of WiVit platform deployment. For each pair of transceivers, when the human is not still, besides the target's moving speed, the change speed of the human reflected path is also related to the target's position. With multiple pairs of transceivers, we can accurately estimate in which area the target is staying. To summarize, WiVit makes the following contributions:

(1) To the best of our knowledge, this is the first platform that is able to monitor the human vitality information with commodity Wi-Fi devices without offline training and calibration. Meanwhile, our platform does not require any dedicated sensor deployment nor require the target to wear any device on the body.
(2) By removing the effect of static multipath and only retrieving the target reflected path change speed from the phase change of Channel State Information (CSI), WiVit could accurately identify when the target is performing activities (even small scale in-place activities such as waving hands) without training or calibration.
(3) Based on the relationship between the target reflected path change speed and the target's position, our platform could detect in which area the target is staying inside a multi-room environment in real-time.
(4) We carry out extensive experiments in three typical indoor environments to evaluate the performance of WiVit. Experiments show that WiVit could detect the target activities at higher precision than 98% in all three environments without any human intervention, demonstrating the robustness of WiVit against environmental changes. Meanwhile, the area detection accuracy is close to 100%.
(5) We further show the potential applications of our platform in practice with two case studies. A short-term activity recognition system is firstly developed upon our WiVit platform, and it reaches an average recognition accuracy of 94.2%. We also take a feasibility study to show that the vitality information captured by the WiVit platform is able to accurately reveal the unique characteristic of each long-term daily activity. We believe that our platform can be utilized to provide valuable datasets to extract high-level semantics of one's different daily life facets, such as living habits, emotions and physical conditions.

The rest of this paper is organized as follows. Section 2 gives an overview of our platform design. Section 3 and 4 introduce the detailed design of our platform. Section 5 presents the implementation and experimental setup. Section 6 shows the evaluation results. Section 7 shows the potential applications of our platform. Section 8 discusses the limitations and future directions. Section 9 introduces the related work followed by a conclusion in Section 10.

2 PLATFORM OVERVIEW

WiVit is a device-free and non-intrusive human vitality monitoring platform. The platform employs existing Wi-Fi transceiver pairs to divide the sensing space into several sub-areas according to the geographical layout of the home environment. Figure 1 shows an example of such a division. The direct path of each pair of transceivers

is the boundary of two adjacent areas. WiVit only leverages CSI samples available at commodity Wi-Fi devices for human vitality monitoring and does not require any human intervention. Following is a brief description of the key steps of WiVit platform:

(1) The first step is to detect whether the target is still or non-still. If the target is non-still, WiVit will record the path change speed spectrum and calculate the path change speed of the target reflected path.
(2) When the target is non-still, WiVit will calculate the approximate human speed and determine whether the target is walking or just performing in-place activities. Then, WiVit will detect which area the target is staying in based on the estimated target reflected path change speeds.
(3) WiVit records the human target's current status (still or non-still), area ID, approximate human speed and the path change speed spectrum at each receiver. Based on these information, we can enable human sensing applications on the platform such as activity recognition.

In the next two sections, we will present in detail how we detect human activity and obtain area information without training or calibration.

3 HUMAN ACTIVITY DETECTION

In this section, we present in detail how WiVit could detect human activities based on the CSI information without training or calibration. We first introduce the relationship between human activities and the Wi-Fi signal. Specifically, the impact of human activities on Wi-Fi signals could be reflected as the CSI phase change in time domain. Then, we describe how to retrieve this phase change from the raw CSI readings aiming for human activity detection. At last, we carry out an empirical study to show that our human activity detection method can work without any human intervention.

3.1 The Influence of Human Activity on Wi-Fi CSI

The CSI reading of Wi-Fi signal is composed of both the amplitude attenuation and phase information in each subcarrier caused by signal propagation from the transmitter to the receiver. For only one path signal, the CSI of the signal at time t_0 can be represented as $x(f, t_0) = A_0 e^{-j2\pi f \tau_0}$, where A_0 is the attenuation of the signal, f is the carrier frequency and τ_0 is the time delay due to propagation. In a typical indoor environment, however, the signal does not only propagate along the direct path but also gets reflected by other objects such as walls, furniture and of course the human target. Thus, the CSI at the receiver side is the superposition of all path signals:

$$x(f, t_0) = \sum_{i=1}^{L} x_i(f, t_0) = \sum_{i=1}^{L} A_i e^{-j2\pi f \tau_i} \tag{1}$$

where L is the number of paths, A_i is the attenuation, τ_i is the propagation delay and $x_i(f, t_0)$ is the CSI of the i^{th} path signal at time t_0. Among all path signals received, those signals that are only reflected from static objects, such as furniture and walls, are defined as *static path signals*, and other signals, which are reflected by a non-still target, are defined as *dynamic path signals*. Note that static path signals include both the direct path signal and signals reflected from walls and furniture which do not change during the process of human activities. When the human target performs activities in the environment, including walking and in-place activities such as waving the hands, the path length of the human reflected signal changes as shown in Figure 2. Suppose the path length change speed is v_{path}, after a short time period t, the CSI of the human reflected signal becomes:

$$x_h(f, t_0 + t) = A_h e^{-j2\pi f(\tau_h + \frac{v_{path} t}{c})} = x_h(f, t_0) e^{-j2\pi f \frac{v_{path} t}{c}} \tag{2}$$

where A_h is the attenuation and τ_h is the propagation delay of the human reflected signal at time t_0, c is the propagation speed of the Wi-Fi signal in the air. For a short time period t, the attenuation change is negligible.

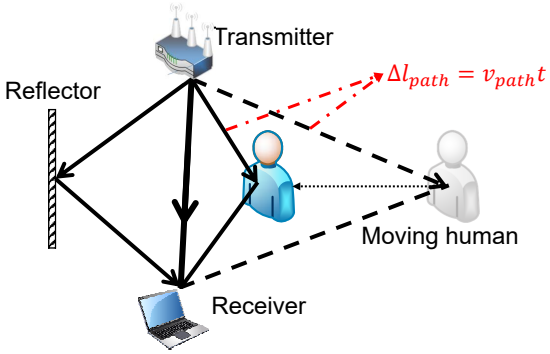
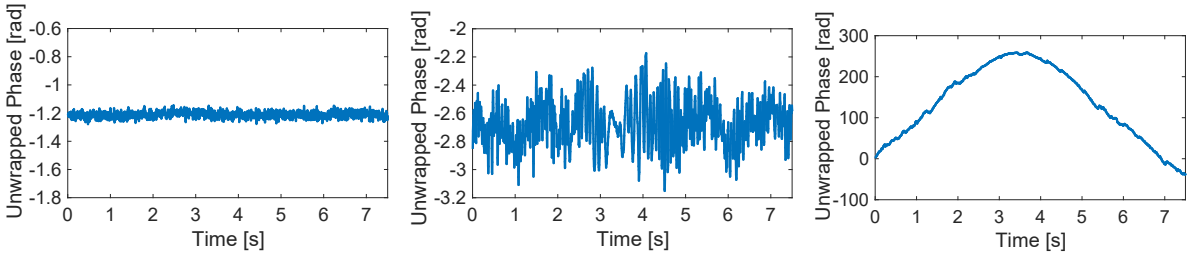


Fig. 2. The human reflected path length will change when the human is non-still.

The path length change is $\Delta l_{path} = v_{path}t$ and thus the propagation delay change is $\Delta\tau = \frac{v_{path}t}{c}$. In the real world with multipath, we can rewrite Equation 1 as:

$$x(f, t_0 + t) = \sum_{i=1}^{L} x_i(f, t_0)e^{-j2\pi f \frac{v_i t}{c}} \tag{3}$$

where v_i is the length change speed of the i^{th} path. If the i^{th} path signal is reflected from a static object, v_i is zero. Here, we conduct a benchmark experiment to show the effect of a moving target on phase change. Figure 3(a) shows the unwrapped Wi-Fi CSI phase[1] when there is no moving human in the environment. Obviously, the phase do not change with time because there is no dynamic path signal. Then, we let a volunteer walk towards the direct path of the transceiver pair and then move away from the direct path. The target reflected path length is expected to decrease first and then increase. Figure 3(b) shows the unwrapped Wi-Fi CSI phase during the process of human movement. Figure 3(c) shows the unwrapped phase change after removing the static path signals. It reflects the phase change of the target reflected path signal. Clearly, the phase of the target reflected path signal changes in accordance with the target movement as expected.



(a) Wi-Fi CSI phase in the static environment (b) Wi-Fi CSI phase when a human is moving (c) Wi-Fi CSI phase of target reflected signal

Fig. 3. (a) shows the unwrapped Wi-Fi CSI phase of a pair of transceivers when the environment is static; (b) shows the unwrapped Wi-Fi CSI phase when a person moves towards the direct path of the transceiver pair and then moves away along the midperpendicular; (c) shows the unwrapped Wi-Fi CSI phase after removing static path signals as described in Section 3.2, which reflects the phase change of the target reflected path signal.

[1]This is the phase after removing the random phase offset as described in Section 3.2.

3.2 Human Activity Detection on Commodity Wi-Fi Devices

Consider the situation when there is only one human reflected path signal at first. When the human target is non-still, the path change speed v_{path} can be considered as a constant in a short time period. Based on Equation 2, the phase change of the CSI reading is also a constant. Other path signals, which are not reflected by the human target do not change with time and will not induce any CSI phase change. Thus, if we can detect such a constant CSI phase change, we know that there exists a dynamic human reflected path signal, and we can further detect the human activity.

Assume we receive M CSI samples in a short time window with each sample timestamped at a microsecond-level precision[2]. If we consider a single path signal, the phase difference between the i^{th} CSI sample and the first sample is $e^{-j2\pi f \frac{v\Delta t_i}{c}}$, where Δt_i is the sampling interval between these two samples and v is the path change speed. The phase differences of the M samples compared to the first CSI sample can thus be expressed as:

$$\vec{a}(v) = [1, e^{-j2\pi f \frac{v\Delta t_2}{c}}, e^{-j2\pi f \frac{v\Delta t_3}{c}}, \ldots, e^{-j2\pi f \frac{v\Delta t_M}{c}}]^\top \tag{4}$$

which is termed as *speed vector*. If the path change speed v is non-zero, the path signal is a dynamic path signal and $\vec{a}(v)$ is the corresponding constant CSI phase change across time. If the path is a static path, v is zero and there is no CSI phase change. With this speed vector, based on Equation 3, the M CSI samples can be represented as:

$$\begin{aligned}
\mathbf{X}(f) &= [x(f,t_0), x(f,t_0+\Delta t_2), \ldots, x(f,t_0+\Delta t_M)]^\top \\
&= \sum_{i=1}^{L} \vec{a}(v_i) x_i(f,t_0) + N(f) \\
&= [\vec{a}(v_1), \vec{a}(v_2), \ldots, \vec{a}(v_L)][x_1(f,t_0), x_2(f,t_0), \ldots, x_L(f,t_0)]^\top + N(f) \\
&= \mathbf{A}\mathbf{S}(f) + \mathbf{N}(f)
\end{aligned} \tag{5}$$

where $N(f)$ is the noise matrix, $\vec{a}(v_i)$ is the speed vector of the i^{th} path signal, $\mathbf{A} = [\vec{a}(v_1), \ldots, \vec{a}(v_L)]$ is the *speed matrix* and $\mathbf{S}(f) = [x_1(f,t_0), \ldots, x_L(f,t_0)]^\top$ is the *signal matrix*. With M CSI samples, if we can calculate the vector $\vec{a}(v)$ for each path signal and detect a non-zero path change speed, we know the human is non-still (either moving or performing some activities).

To estimate the vector $\vec{a}(v)$ for each path signal, we apply the MUSIC algorithm [32] on Equation 5. The basic idea of MUSIC algorithm is eigenstructure analysis of the $M \times M$ correlation matrix $\mathbf{R_x}$ of the received M CSI samples. From Equation 5, we obtain $\mathbf{R_X}$ as:

$$\begin{aligned}
\mathbf{R_X} &= \mathbb{E}[\mathbf{X}\mathbf{X}^{\mathrm{H}}] \\
&= \mathbf{A}\mathbb{E}[\mathbf{S}\mathbf{S}^{\mathrm{H}}]\mathbf{A}^{\mathrm{H}} + \mathbb{E}[\mathbf{N}\mathbf{N}^{\mathrm{H}}] \\
&= \mathbf{A}\mathbf{R_S}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}
\end{aligned} \tag{6}$$

where $\mathbf{R_S}$ is the correlation matrix of the signal matrix $\mathbf{S}$, $\mathbf{I}$ is an identity matrix and σ^2 is the variance of noise. The correlation matrix $\mathbf{R_X}$ has M eigenvalues. The eigenvectors corresponding to the smallest $M-L$ eigenvalues construct a noise vector subspace $\mathbf{E_N} = [\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_{M-L}]$, and the other L eigenvectors construct a signal subspace $\mathbf{E_S} = [\vec{e}_{M-L+1}, \vec{e}_{M-L+2}, \ldots, \vec{e}_M]$. The signal and the noise subspaces are orthogonal. If a path signal exists, the corresponding speed vector is orthogonal with the noise subspaces. Thus, the speed spectrum function is expressed as:

$$P(v)_{MUSIC} = \frac{1}{\vec{a}^{\mathrm{H}}(v)\mathbf{E_N}\mathbf{E_N}^{\mathrm{H}}\vec{a}(v)} \tag{7}$$

[2]This is supported by commodity Wi-Fi cards such as Intel 5300.

in which a peak corresponds to a signal and occurs at the corresponding path change speed.

However, to obtain the path change speed spectrum with commodity Wi-Fi devices, one more challenge needs to be addressed. The commodity Wi-Fi transceivers are not tightly synchronized with each other. So a time-variant random phase offset $e^{-j\theta_{offset}(t)}$ exists between two adjacent CSI samples, which distorts the CSI phase change in time domain and prevents us from getting the right speed spectrum estimation, as shown in Equation 8:

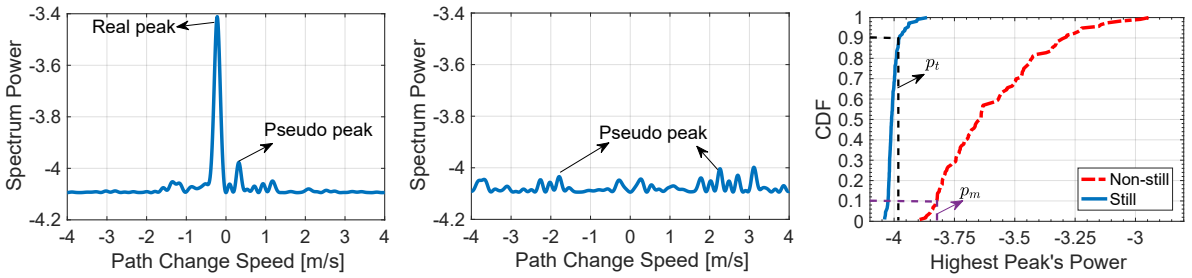
$$x(f, t_0 + t) = e^{-j\theta_{offset}(t)} \sum_{i=1}^{L} x_i(f, t_0) e^{-j2\pi f \frac{v_i t}{c}} \tag{8}$$

Fortunately, we find that the time-variant random phase offsets are the same across different antennas on a same Wi-Fi card because they share the same RF oscillator. Therefore, we can apply conjugate multiplication between the CSI readings from two antennas to remove this time-variant phase offset. Meanwhile, we also remove the static component and adjust the weight of the antennas as described in [21] to ensure accurate spectrum estimation.

After this processing, we could obtain the path change speed spectrum based on Equation 7, as shown in Figure 4(a) and Figure 4(b). Since we have removed all static path signals, there is no real peak on the spectrum when the human is still. The height (i.e. spectrum power) of each peak represents the degree of orthogonality between the signal subspace and noise subspace, and can be considered as the probability of the existence of a path signal. The spectrum power of a real peak, which corresponds to a real path signal, is much higher than those small pseudo peaks. We carry out benchmark experiments to demonstrate this. We ask a person to hold his breath and keep still for 5s to ensure all path signals are static. Then, we ask the person to keep non-still for another 5s. Figure 4(c) shows the cumulative distribution function (CDF) of the absolute power of the highest peak on the spectrum when the target is still and non-still. We take the 90^{th} percentile of power (p_t) when the person is still to represent the power level of the pseudo peaks. Meanwhile, we take the 10^{th} percentile of power (p_m) when the person is non-still to represent the power level of the real peaks. We can see that there is a clear gap between the two values (p_t and p_m) so a simple threshold can be applied to easily identify the real peaks. WiVit platform employs the p_t as the threshold for human activity detection.

3.3 Robustness of Human Activity Detection

Human activity can thus be detected based on the spectrum power. If the power of the highest peak is larger than the threshold p_t, then the human target is considered to be non-still. The path change speed, at where the

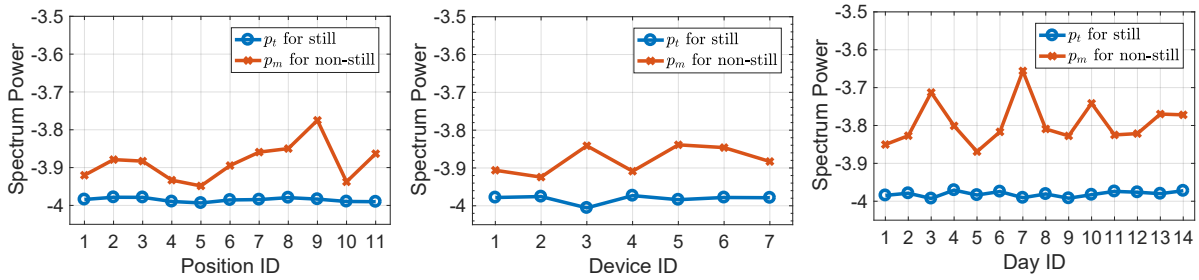


(a) Speed spectrum when the target is non-still (b) Speed spectrum when the target is still (c) CDF of highest peak's power

Fig. 4. (a) is a path change speed spectrum when the target is non-still; (b) is a path change speed spectrum when the target is still; (c) is the CDF of highest peak's power when the target is still and non-still.

highest peak occurs, is the target reflected path change speed. As the constant phase change caused by human movement can not be induced by environmental changes and noise, the spectrum power obtained is not affected by environmental changes or noise either. Here, we carry out an empirical study to show the robustness of the spectrum power.

We place a pair of Wi-Fi transceivers at different locations in two different environments: one multi-room home and one office room. In both environments, there are furniture and electronic appliances which can generate multipath. Thus, when Wi-Fi transceivers are placed at different locations, the transceivers have significantly different multipath propagations. For each location, we let a human target hold his breath and keep still for 5s and record the power of the highest peak to get the p_t value. We also let the same human target keep non-still for 5s and record the power of the highest peak to get the p_m value. Figure 5(a) shows the p_t and p_m values at different locations in these two environments. We can see that there is a clear gap between p_t and p_m curves. Meanwhile, we can see that the threshold p_t is stable across different transceiver locations and different environments. This means that we do not need to adjust the p_t value even when there are environmental changes.



(a) Impact of different environments (b) Impact of different devices (c) Impact of different days

Fig. 5. (a) shows the power level of real and pseudo peak in different multipath environments (i.e. the positions of devices). The position 1-7 are in the smart home and position 8-11 are in the office room; (b) shows the power level of real and pseudo peak on different devices; (c) shows the power level of real and pseudo peak in different days.

We further evaluate the robustness of the threshold p_t against hardware diversity. We put different Wi-Fi devices at the same position as the receiver and repeat the experiments mentioned above. Figure 5(b) shows the p_t and p_m values for different devices and we can clearly see that the threshold p_t is stable and the gap between p_t and p_m is also clear. We also evaluate the stability of p_t for a long term. We keep WiVit running continuously for 14 days and record the spectrum power at the same time. The results are shown in Figure 5(c). The threshold p_t value is still stable.

So we can conclude that the spectrum power threshold p_t for human activity detection is stable across different environments, devices and time periods. Thus, we can detect human activities without any human intervention. For the human activity detection module of WiVit, the detection threshold is initialized with a default value[3]. Meanwhile, in order to improve the detection accuracy, the platform will update the threshold value automatically when it detects still status of a target for longer than 5s. The new threshold is updated with the new p_t obtained in the latest 5s when the target is still.

[3]The default value is set as -3.97.

4 AREA DETECTION

In this section, we describe in detail how WiVit identifies which area the target is located without any offline training or calibration. We first introduce the relationship between the target reflected path change speed and the target's position. Then, we propose a method to directly calculate which area the target is located based on the path change speeds. At last, we introduce how WiVit could segment when the target is walking in a continuous activity sequence for accurate area detection.

4.1 The Relationship between Human Target's Position and Path Change Speed

For a pair of Wi-Fi transceivers, we can assume the human target is located at an ellipse with foci at the transmitter and receiver, as shown in Figure 6. The human speed can be decomposed into two parts: the normal speed and the tangent speed. The tangent speed is parallel to the ellipse and does not change the length of human reflected path. The normal speed is perpendicular to the ellipse and will cause changes in path length. When the human target is at different positions but with the same human speed, the normal speed will be different and thus will cause a different path length change speed. If we know the positions of the Wi-Fi transceivers ($\vec{P}_{tx}$ and $\vec{P}_{rx}$), the relationship between the target's position ($\vec{P}_h$) and the path length change speed can be expressed as:

$$v_{path} = \vec{v}_h \cdot \frac{\vec{P}_h - \vec{P}_{tx}}{\|\vec{P}_h - \vec{P}_{tx}\|} + \vec{v}_h \cdot \frac{\vec{P}_h - \vec{P}_{rx}}{\|\vec{P}_h - \vec{P}_{rx}\|} \tag{9}$$

where $\vec{v}_h$ is the human speed.

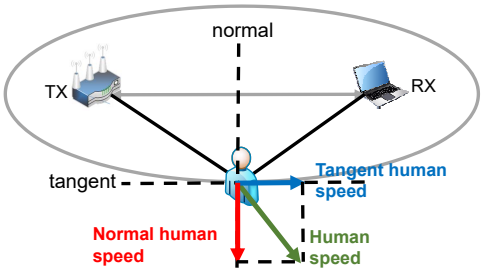


Fig. 6. The relationship between the human speed and the human reflected path change speed.

4.2 Speed-based Area Detection

Based on Equation 9, if we know the human speed, with more than one pair of Wi-Fi transceivers whose positions are known, we can calculate the human target's position. However, in reality, we also do not know the human speed. To solve this problem, based on the geographical layout of the home environment, we utilize multiple pairs of Wi-Fi transceivers to divide the monitoring space into several areas. The direct path between a pair of transceivers is the boundary of two areas. Figure 1 shows a deployment example of WiVit platform. Based on the target reflected path change speed estimated on these receivers, WiVit can accurately identify in which area the target is staying.

The basic idea of area detection is that, we first assume the target is staying in a particular area say the z^{th} area, we can then calculate the approximate human speed based on the path change speeds at the two receivers on the boundary of the area, such as the receiver 1 and 2 for area 1 in Figure 1. The target movement also induces non-zero path change speeds at other receivers. With the human speed estimated by assuming the target is located in area z, we can calculate the expected target reflected path change speeds at other receivers. On

the other hand, the actual target reflected path change speeds can also be calculated with measured CSI phase changes. We now obtain two copies of path change speeds at each other receiver and we compare the two copies of speeds. If the initial assumption that the target is located in the z^{th} area is correct, these two copies of path change speeds will match very well. Otherwise, our assumption is not correct and we continue to assume the human target is located in another area and repeat the process above. The detailed steps are shown as below:

(1) **Approximate human speed estimation in each area.** For a specific area, we assume the target stays at the center of the area. Based on Equation 9 and the target reflected path change speeds at two receivers on the boundary of the area, we can calculate the human speed. However, since the target can be at any location in the area and may not be at the center, this estimated speed is an approximate speed.
(2) **Current area status estimation.** Assuming the target is staying in the z^{th} area, we obtain the approximate human speed in step 1. From the obtained approximate speed, we can calculate the expected target reflected path change speeds at other receivers excluding the two receivers selected in step 1. On the other hand, we can also employ the measured phase changes at other receivers to directly get the path change speeds. If the two copies of path change speeds match well, we can conclude our assumption is correct and the target is located in the z^{th} area. We define the possibility the target is located in the z^{th} area as $Prob_z$ and calculate its value as below: :

$$Prob_z = \frac{K}{\sum_{i \in D_z} (v_i - \hat{v}_{z,i})^2} \tag{10}$$

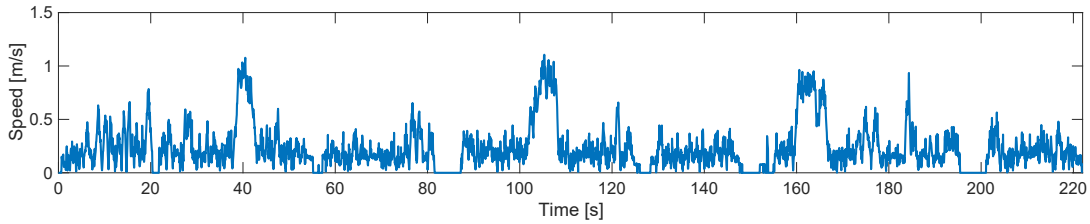
where D_z is the set of receivers which are not used to calculate the approximate human speed in the z^{th} area; K is the number of receivers in set D_z; v_i is the target reflected path change speed directly obtained from the CSI phase measurement at the i^{th} receiver and $\hat{v}_{z,i}$ is the expected change speed calculated with the approximate human speed estimated in the z^{th} area. In this way, we can get the probability for each area and then choose the one with the highest probability as the human target's staying area.
(3) **Enhance the accuracy of area detection.** When the target moves to another area, he/she will cross over the direct path, which is the boundary of two adjacent areas. Throughout this process, the target moves towards the direct path first, crosses over and then moves away. Thus, the target reflected path change speed at the receiver will be negative first, change to zero, and finally become positive. To improve the area detection accuracy, WiVit only updates the human area status when such a change pattern of the path change speed is detected at the boundary receiver.

With the above three steps, WiVit can detect which area the target is staying in without any offline training or calibration.

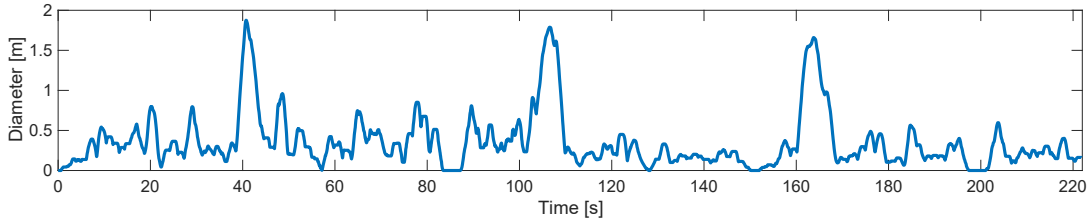
4.3 Walk Detection for Accurate Area Detection

However, when the target is not walking but only performing in-place activities, the path change speed caused by the target's movement is small and irregular. Meanwhile, the target will not move to another area during performing in-place activities. Thus, the WiVit platform detects which area the target is staying in based on the path change speeds only when the target is walking.

To distinguish in-place activities and walking, the basic idea is that when the target is performing in-place activities, the target's position only varies in a small range. On the other hand, the target's position has much larger change during walking. Based on the approximate human speed, WiVit can calculate the approximate human displacement and obtain the target's position sequence. Within a time window, WiVit can obtain many position estimations and employ a circle to cover all these positions. The diameter of this circle indicates the range of the position changes within this time window. We can then set a threshold. If the diameter is larger than the threshold, we can identify the target is walking but not performing in-place activities. For WiVit platform, the window size is set as $2s$ and the diameter threshold is $0.8m$. We ask a volunteer to walk for a short period



(a) Approximate human speed



(b) Diameter of human position change in 2s

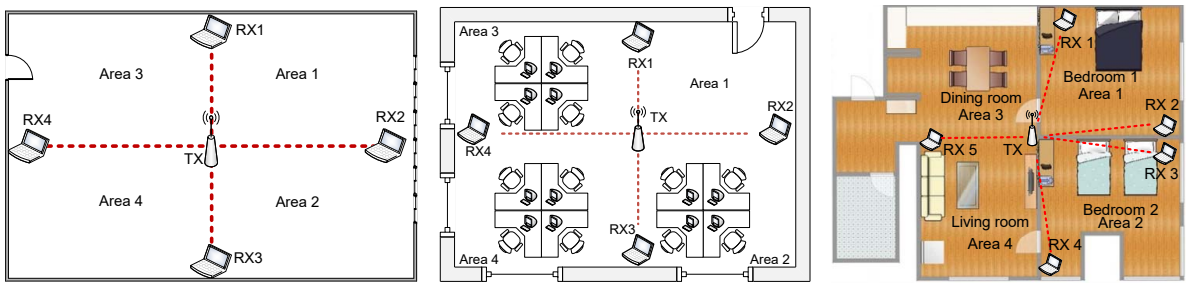
Fig. 7. Comparison between the approximate human speed and the diameter of human position change: (a) shows the approximate human speed with three times of human walking; (b) shows the corresponding diameter of human position change in each 2s time window and the moving step is 0.05s.

of time for three times. Between two adjacent walks, the volunteer either performs in-place activities freely or keeps still for a while. We present the results in Figure 7 and we can see that the diameter plot is much smoother than human speed plot. It will be more accurate to identify walking according to the diameter of the human displacement. Thus, we adopt human displacement for walk detection.

5 EXPERIMENTAL SETUP

We employ miniPCs equipped with cheap off-the-shelf Intel 5300 Wi-Fi cards as the transmitter and receivers. Each receiver is attached with two antennas. The CSI tool developed by Halperin [10] is installed on these miniPCs to collect CSI samples of each received packet. The sampling rate of CSI for WiVit platform is 200 Hz. For each path change speed estimation, we employ CSI samples collected over a period of 0.3s. Each receiver calculates the path change speed spectrum with the received CSI samples and sends the spectrum to a server for human vitality computation in real time. WiVit platform can be hosted on any channel on the 2.4 and 5 GHz bands. To avoid interference from ongoing data communication, we employ an unused 20 MHz channel on the 5 GHz band. Also any type of Wi-Fi packet including beacons can be employed for our platform. Thus, WiVit platform has a minimum impact on the existing Wi-Fi data communication. When we deploy our WiVit platform, we measure the positions of Wi-Fi transceivers carefully with a laser range meter.

To evaluate the performance of WiVit platform, we conduct experiments in three typical indoor environments: a large empty room, a large office room and a real multi-room smart home, with dimensions labeled as shown in Figure 8. In the office room and multi-room smart home, there are many furniture and electronic appliances so rich multipath exists. For activity detection, WiVit only needs one receiver to capture the path change speed incurred by the target's activities. Thus, in each room/area, one receiver is enough and WiVit could detect when the target is non-still accurately. For area detection, WiVit needs at least two receivers in each room/area to capture the complex path change speed incurred by the target's movements. Therefore, in the empty room and



(a) Empty room (b) Office room (c) Smart home

Fig. 8. Experiment environments: (a) is an empty room and each area is $3.5m \times 4m$; (b) is an office room, in which area 1 is $3.5m \times 4m$, area 2 is $4m \times 4m$, area 3 is $3.5m \times 5m$ and area 4 is $4 \times 5m$; (c) is a smart home, in which the bedroom 1 is $4.2m \times 4.9m$, the bedroom 2 is $5m \times 4.9m$, the dining room is $4.2m \times 4m$ and the living room is $5m \times 4m$.

Table 1. Information about different volunteers

Volunteer ID	Gender	Age	Height [cm]	Weight [kg]
1	Male	29	183	102
2	Male	27	172	78
3	Female	24	158	54
4	Female	27	164	52
5	Female	23	170	61

office room, we use 1 transmitter and 4 receivers to divide the sensing space into 4 areas. In the multi-room smart home, we employ 5 receivers to divide the home into 4 areas: two bedrooms, a dining room and a living room. As shown in Figure 8, we place each receiver at the corner of an area to ensure the line-of-sight (LoS) path is the boundary of two adjacent areas. To evaluate the performance of WiVit, we ask 5 volunteers who are graduate students in our lab to be our experimental subject. Table 1 shows the basic information of these volunteers. In order to evaluate the robustness of WiVit platform for long-term sensing, we keep the platform running for 14 days continuously in the smart home environment. For basic vitality status sensing evaluation, each participant could freely walk, move across areas, keep still and perform in-place activities in all the four areas. For each experiment, we only require the participant to keep still, walk around and perform in-place activities in each area at least once. We do not apply any other restriction on the participants. In the smart home environment, the total recorded number of activities is 3427, including 2015 walkings (756 area changes). In the office room, the total number of activities is 410, including 240 walkings (120 area changes). In the empty room, the total number of activities is 373, including 204 walkings (100 area changes). In each environment, we employ cameras to record all the targets' activities as the ground truth. Figure 9 shows the graphical user interface of WiVit platform which shows both the ground truth and the human status sensed by the platform.

6 BASIC VITALITY STATUS SENSING

In this section, we evaluate the performance of WiVit for human vitality status sensing, including activity detection, walk detection and area detection. For area detection, we plot the area ID sequence detected during the monitoring process and compare it with the ground truth. For activity detection and walk detection, we employ two commonly used metrics *precision* and *false negative rate (FNR)* to show the performance. We formally define

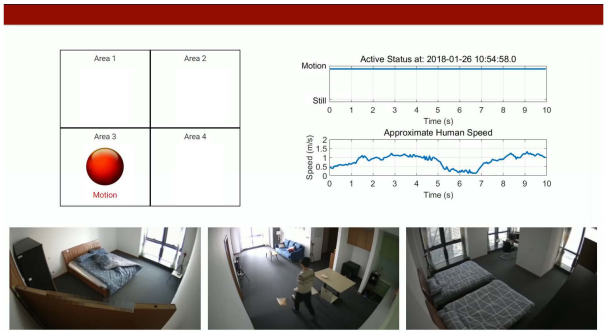


Fig. 9. Graphical user interface of WiVit platform.

the two metrics as below:

$$precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{11}$$

$$FNR = \frac{FalseNegative}{TruePositive + FalseNegative} \tag{12}$$

6.1 Human Activity Detection

6.1.1 Overall Performance. In the multi-room smart home environment, we keep WiVit platform running for 14 days and evaluate the performance of long-term human activity detection. When the target is moving, the status is non-still. If the target is not moving but performing in-place activities such as waving the hands, the ground truth is also non-still; otherwise the ground truth status is still. Table 2 shows the performance of activity detection of WiVit platform. Without any human intervention, the precision of human activity detection is about 98% and FNR is about 1%. These results demonstrate that our human activity detection module is robust for long-term human sensing.

Table 2. Performance of human activity detection in different days

Day ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Precision [%]	98.8	97.6	98.9	98.8	97.8	98.1	98.5	99.1	97.6	97.8	98.3	99.3	97.2	97.7
FNR [%]	1.15	1.75	1.38	0.98	0.88	0.91	0.89	1.44	1.32	1.69	1.04	0.89	1.28	1.35

6.1.2 Impact of Human Diversity. To evaluate whether WiVit has a consistent performance for different human targets, we collect activity data of 5 volunteers in the multi-room smart home. Table 3(a) shows the precision and FNR for different participants. We can see that, for different participants, WiVit could achieve activity detection precision higher than 97% and FNR lower than 1.97%.

6.1.3 Impact of Different Environments. We also evaluate the platform performance in different indoor environments. Table 3(b) shows the precision and FNR of activity detection in an empty room, an office room and a multi-room smart home, respectively. Although there are much more multipath in the office and multi-room smart home than in the empty room, the performances of activity detection are similar. The results demonstrate the robustness of our platform against environmental changes.

Table 3. Impact of different participants and environments for human activity detection

(a) Impact of different participants

Participant	1	2	3	4	5
Precision [%]	97.9	97	98.1	97.4	97.6
FNR [%]	0.97	1.57	1.97	1.24	1.49

(b) Impact of different environments

Environment	Empty room	Office room	Smart home
Precision [%]	98.9	97.8	98.3
FNR [%]	1.49	1.34	1.28

Table 4. Performance of human activity detection for different daily activities

Daily activity	Sleep	Eat	Watch TV	Clean
Precision [%]	99.1	97.4	98.2	99.6
FNR [%]	0.88	1.31	1.23	0.18

6.1.4 Impact of Different Activities of Daily Living. Moreover, in order to evaluate the performance of different activities of daily living, we ask the volunteers to perform 4 types of activities in the smart home: sleeping, eating, watching TV and cleaning the floor. Note that during the process of a daily activity, the participant does not keep non-still all the time, but switches between still and non-still alternatively. For example, when the participant is sleeping, he is still most of the time but non-still when he adjusts the posture. As shown in Table 4, for different daily activities, WiVit could always achieve a consistent good performance for activity detection.

6.2 Human Walk Detection

Table 5 shows the overall performance of walk detection when the human is non-still. The precision of walk detection is around 96% and the FNR is 5%. In real life, some non-walking activities can be detected as walking. For example, when the participant sits down, the torso has a large displacement in a short time. This displacement is similar to the torso displacement when the human is walking. Moreover, some miss detections may also occur when the participant is walking. This is because when the participant turns his/her body to change the walking direction, the absolute displacement is very small so WiVit may identify this movement as in-place activities rather than walking.

We further evaluate the performance of walk detection with different participants in different environments. As shown in Table 6(a) and Table 6(b), WiVit could achieve consistent high precision of walk detection for different participants and in different environments.

6.3 Human Area Detection

To evaluate the area detection performance, we compare the area detection results with the ground truth. The ground truth is obtained by using cameras. In the empty room, each area has the same size of $3.5m \times 4m$. In the office room, area 1 is $3.5m \times 4m$, area 2 is $4m \times 4m$, area 3 is $3.5m \times 5m$ and area 4 is $4m \times 5m$. The total space of the smart home is $82m^2$ and is divided into 4 rooms. The size of bedroom 1 is $4.2m \times 4.9m$, the bedroom 2 is $5m \times 4.9m$, the dining room is $4.2m \times 4m$ and the living room is $5m \times 4m$.

Table 5. Performance of human walk detection in different days

Day ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Precision [%]	96.3	96.7	94.4	95.6	96.4	97.4	96.5	95.6	94.8	97.8	96.1	97.5	94.5	95.9
FNR [%]	4.45	4.96	6.39	5.65	5.79	6.46	4.62	5.12	5.63	4.96	4.18	6.98	5.24	5.36

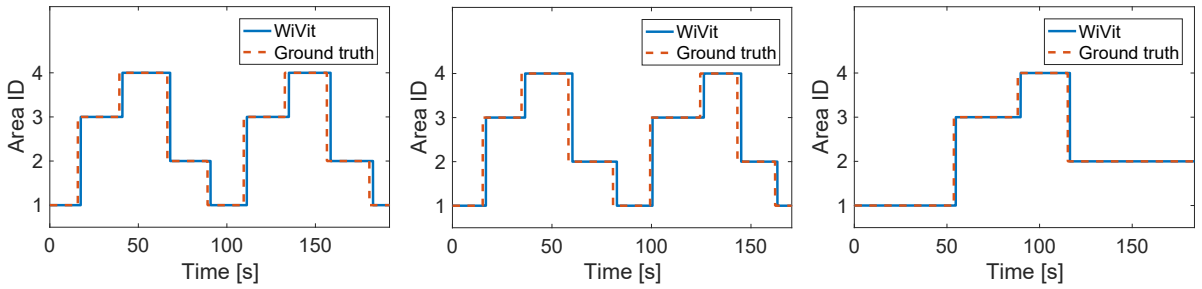
Table 6. Impact of different participants and environments for human walk detection

(a) Impact of different participants

Participant	1	2	3	4	5
Precision [%]	96.9	95.6	94.4	97.2	96.8
FNR [%]	5.44	4.67	5.21	4.36	5.89

(b) Impact of different environments

Environment	Empty room	Office room	Smart home
Precision [%]	94.9	95.5	96.12
FNR [%]	4.85	5.07	5.36



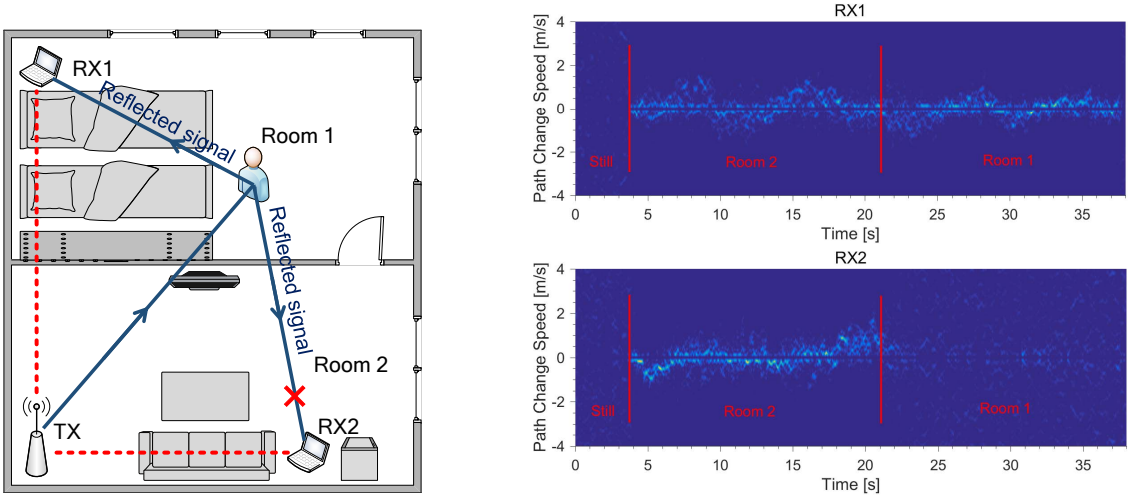
(a) Empty room (b) Office room (c) Smart home

Fig. 10. Performance of area detection in different environments.

Figure 10 shows the accuracy of area detection in different environments. In each environment, WiVit could correctly identify which area the participant stays in and when the participant moves to another area. Note that, compared with the ground truth, the time points when the human changes area estimated by WiVit have a short latency. This is because our platform requires buffering several CSI samples before estimating the path change speed.

6.3.1 Possibility of Reducing the Number of Devices for Area Detection. In this paper, we divide the sensing space into 4 areas and each area is bounded by two transceiver pairs (streams). In real world, some houses may have more rooms and a more complex floorplan with corridors/turns. In these houses, to capture the target information in the corridor/turn and each room, the easiest approach is to deploy more Wi-Fi devices to make sure each room/corridor/turn is bounded by two streams. However, this approach requires a high deployment density. We further observe that, even sometimes when two rooms are covered by only two streams, based on the signal strengths at different receivers, it is still possible for us to decide which room the target is staying in without deploying more Wi-Fi devices. The same concept can also be applied to detect whether the target is in the corridor/turn.

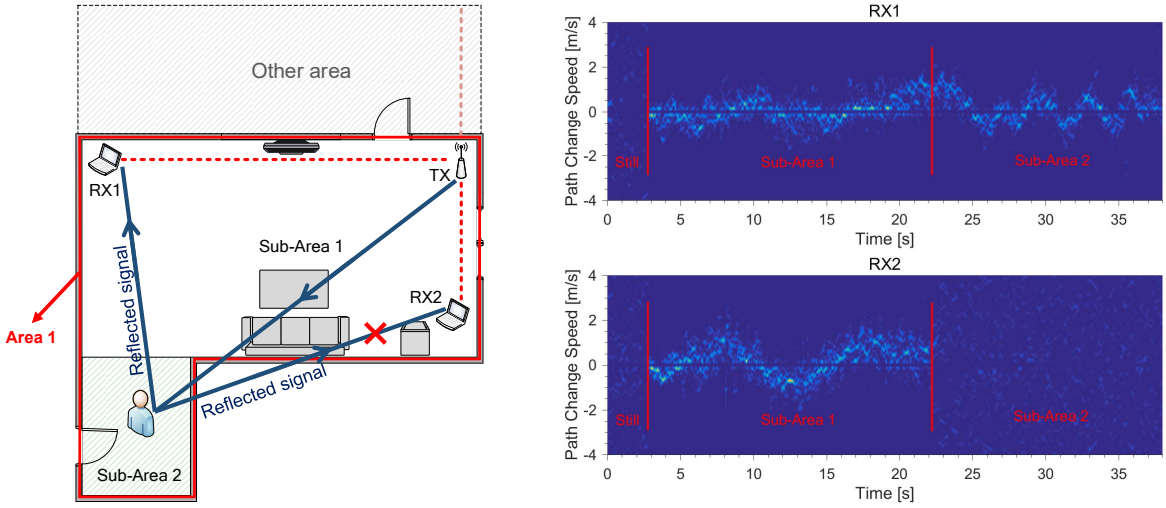
As shown in Figure 11(a), we use two transceiver pairs (streams) to bound an area which contains two rooms. We deploy the transmitter and receivers at corners of the area and there is no receiver placed at the boundary of the two rooms. We ask a volunteer to keep still in room 2 for a while at first and then move around for a moment before move into room 1. Figure 11(b) shows the path change speed spectrums at the two receivers during the process. When the target keeps still, there is no dynamic path signal on the spectrums. When the target stays in room 2 and keeps moving, both of two receivers could detect the human movement on the spectrum. But when the target stays in room 1 and keeps moving, only the receiver 1 (RX1) could detect the human movement on the spectrum. This is because the target reflected signal to receiver 2 (RX2) is blocked twice by the wall so that the signal is too weak to be sensed. However, the target reflected signal to receiver 1 (RX1) is only blocked once by



(a) The layout of an area contains two rooms

(b) The path change speed spectrums of two receivers

Fig. 11. (a) shows the layout of an area contains two rooms; (b) shows the path change speed spectrums of two receivers when the target stays in the area.



(a) The layout of an area contains a corridor (the sub-area 2)

(b) The path change speed spectrums of two receivers

Fig. 12. (a) shows the layout of an area contains a corridor (the sub-area 2); (b) shows the path change speed spectrums of two receivers when the target stays in the area.

the wall so that the human movement can still be detected on the spectrum of receiver 1. Thus, based on the sensibility of different receivers, we can decide which room the target is staying in.

As shown in Figure 12(a), we use two transceiver pairs (streams) to bound the area 1, which contains a corridor (the sub-area 2), to show how to deal with complex floorplan with corridors/turns. Figure 12(b) shows the path change speed spectrums of the two receivers when the target stays in the area. When the target keeps still in the sub-area 1, there is no dynamic path signal existing on the spectrums. When the target moves around in the sub-area 1, both of the two receivers could detect the human movement. On the other hand, when the target stays in the sub-area 2 and moves, the receiver 2 (RX2) cannot detect the target's movement on the spectrum. This is because the target reflected signal to receiver 2 is blocked by two walls so the reflected signal is too weak to be sensed. Thus, based on the signal changes at different receivers, we can decide when the target moves to the sub-area 2 (corridor) without deploying any extra Wi-Fi device.

Although we can locate the target in a larger and more complex area without deploying more Wi-Fi devices, we should note that this convenience is achieved at the expense of losing some important sensing information. For example, when the target stays in the sub-area 2 as shown in Figure 12(a), we will lose the speed information about the target at receiver 2. In the future, we plan to employ LoRa which is designed for the next generation IoT connections. The communication range of LoRa can be much larger so 1 to 2 transceiver pairs are able to cover a larger and more complex area.

7 USE CASE STUDY

7.1 Use Case 1: Human Action Recognition

Human activities could be divided into short-term actions and long-term activities of daily living [8]. The term "action" is the physical action of a user and typically lasts for a short duration of time, such as sitting down, falling, etc. In this section, in order to show the potential applications of our platform, we choose 4 basic and commonly seen human actions (walking, running, sitting down, and falling) and develop an activity recognition system on our WiVit platform to recognize them. As described in Section 3, WiVit platform is able to obtain the path change speed spectrum at each receiver. Based on Equation 3, the received signal is the superposition of all path signals at the receiver. Thus, if two parts of human body move at different speeds, there will be two dynamic path signals with different path change speeds on the spectrum. For different activities, the moving speed characteristics of the human body are very different [40]. With the path change speed spectrum, we can build activity model to recognize different activities.

For activity recognition, we use Hidden Markov Model (HMM) to build the activity model for each activity. For an activity sample, we use a sequence of feature vectors to represent the whole process. The feature vector is extracted from the path change speed spectrum at the receiver. We divide the path change speed range $0m/s$ to $4m/s$ into 20 components equally. The power of each component is the sum of spectrum power of all speeds in that component. Thus, for each spectrum estimation, we can get a 20-dimensional feature vector. Our platform outputs a spectrum estimation each $50ms$. Moreover, since an area is bounded by two receivers, for an activity sample, we will get two sequences of feature vectors. Based on the well-known Baum-Welch algorithm [43], we can construct an HMM for each activity with the training samples of that activity. During the training phase, we use both sequences of feature vectors of each activity sample to learn the activity model. During the classification phase, we apply Likelihood Fusion [40] to combine the two sequences of feature vectors of an activity sample for classification. In the smart home environment, we collect 200 samples (50 samplers per area) for each activity. These samples are collected from the 5 volunteers in Table 1, and they serve as the training dataset. In the empty room and office room, respectively, we collect 40 samples (10 samples per area) for each activity from two new volunteers, whose data is not in the training dataset, to evaluate the robustness of our activity model for untrained environments and human targets. Note that, although the activity recognition system requires training the activity model, the required information for action segmentation (when the target is non-still) and feature extraction (path change speed spectrum) can be provided by our platform without training or calibration.

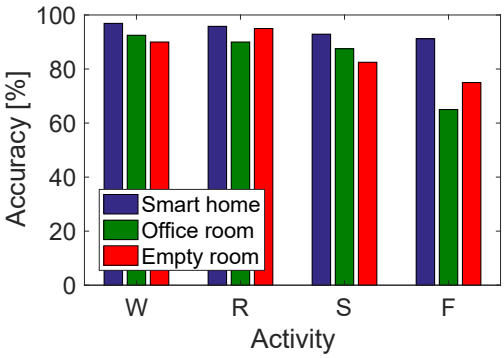


Fig. 13. The recognition accuracy of 4 actions: walking (W), running (R), sitting down (S) and falling (F).

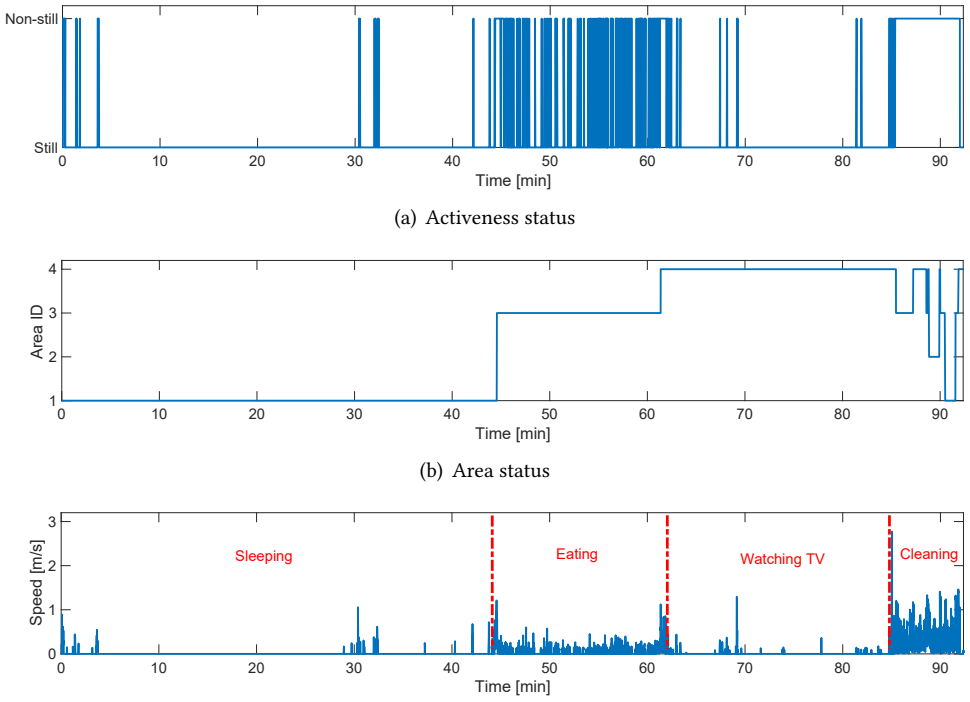
We evaluate the performance of activity recognition in terms of recognition accuracy, which is calculated as the number of correctly recognized activities divided by the total number of activities performed. As shown in Figure 13, in the smart home, the average 10-fold cross-validation accuracy is 94.2% across all activities. In the untrained environments with untrained new targets, the average accuracy is 84.7%.

7.2 Use Case 2: Real Life Human Vitality Sensing

Moreover, we also carry out a feasibility study about monitoring long-term activities of daily living based on our WiVit platform. We ask the 5 volunteers in Table 1 to live in the smart home for several hours everyday. During their staying, each participant is asked to perform 4 types of daily activities freely: sleeping, eating, cleaning the floor and watching TV. Among the 4 types of daily activities, 56% of the time is sleeping, 12% is eating, 24% is watching TV and 8% is cleaning the floor. Then we extract the vitality information captured by WiVit for these daily activities.

7.2.1 Basic Information of Human Vitality. Figure 14(a)-14(c) show the basic human vitality information, including the activeness status, human speed information and the area detection results, obtained from an actual daily life scenario over 90 minutes. In this scenario example, the participant sleeps in bedroom 1 (i.e. area 1) first and then walks to the dining room (i.e. area 3) for a meal. After eating, the participant walks to the living room (i.e. area 4) and watches TV. After watching TV, the participant cleans the floors of all rooms. Based on the basic information shown in Figure 14, we can further obtain the human vitality statistics about daily activities. In this paper, we focus on 3 human vitality statistics: *(i) human active rate*, which is the fraction of time when the participant is non-still; *(ii) area staying rate*, which is the time fraction of the participant stays in each area; and *(iii) average approximate human speed*, which indicates how fast the human body moves when the participant is non-still.

7.2.2 Human Vitality of Different Daily Activities. We first carry out experiments to study the human vitality of different daily activities. Based on the long-term daily activity data collected in the smart home, we calculate the human vitality statistics for different activities. As shown in Figure 15, the human vitality statistics clearly show the differences between different daily activities. For sleeping and watching TV, the human active rate is very low. It is because the human keeps still most of the time when lying on the bed or sitting on the sofa. Only when the participant adjusts the body posture or position during sleeping or sitting, non-still human activity can be detected. However, for sleeping, the human stays in the bedroom (i.e. area 1) while for watching TV, the human is in the living room (i.e. area 4). When the participant is eating, the participant is at the dining room (i.e.



(a) Activeness status

(b) Area status

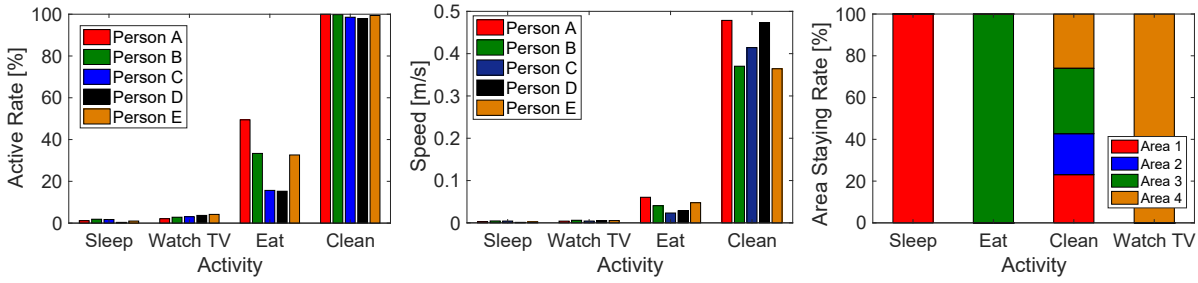
(c) Approximate human speed

Fig. 14. Basic human vitality information: (a) shows the human activeness status of an actual daily life scenario; (b) shows the area status and (c) shows the approximate human speed.

area 3) and the active rate is much higher than sleeping and watching TV as the human needs to continuously move his hands for eating. However, the movement speed is quite low. When the participant is cleaning the floor, the participant has the highest active rate among the four activities. Meanwhile, the average human speed is much higher than other three activities as shown in Figure 15(b). Moreover, for cleaning, the human moves across all four areas.

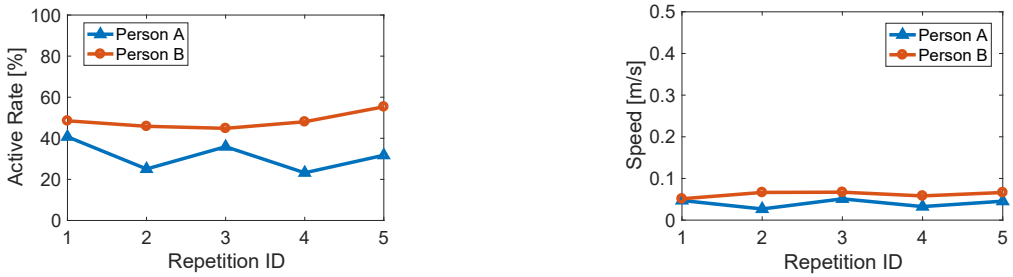
7.2.3 Human Vitality of Different Participants. We also study the captured human vitality of different participants in this section. As the active area is exactly the same for different participants, we only compare the active rate and average human speed across different participants. As shown in Figure 15(a) and 15(b), for most daily activities, different participants have very similar vitality statistics except for eating activity. For eating activity, two participants have lower active rates and speeds than others. This is because that the two female volunteers are gentler than others during eating.

7.2.4 Human Vitality of Different Repetitions. Finally, we would like to study the human vitality of different repetitions for the same activity. In Figure 16, we show that the active rate and average human speed of different repetitions for eating. The data is collected from two volunteers and each performs the eating action for 5 times in different days. Even for the same participant, in different repetitions, the active rate and human speed have slight differences. However, the active rate and human speed of the eating activity are still much higher than sleeping and watching TV and lower than cleaning floors as shown in Figure 15.



(a) Human active rate (b) Average human speed (c) Area staying rate

Fig. 15. Human vitality statistics for different activities and participants.



(a) Human active rate (b) Average human speed

Fig. 16. Human vitality statistics for different repetitions of the eating activity.

8 DISCUSSION

In this work, we design an indoor human vitality sensing platform that is device-free and utilizes only commodity Wi-Fi devices. We hope this platform could provide support for a variety of Wi-Fi based device-free sensing applications. There are several directions to further extend our work, which we discuss below.

Multi-person vitality sensing. Passive or device-free sensing of multiple targets is known to be challenging. When there are multiple persons, if only the target is non-still, our platform could still capture the target's vitality information accurately. If there are more than one non-still person, the signal received at the receiver is a superposition of dynamic reflected signals from different persons. Due to the small Wi-Fi bandwidth (20MHz), distinguishing mixed signals at commodity Wi-Fi devices is still a big challenge and we consider it as one important direction of our future work. We may consider channel hopping to multiple channels and combine information from multiple channels to form a virtual larger channel to address this issue.

More human daily activity dataset. As we need deploy cameras in each room to record the ground truth, which poses severe privacy concerns, WiVit is only deployed in a smart home environment to record daily activities performed by volunteers. In the future, we plan to deploy our WiVit platform in real home environments and record human daily activities for months.

Other moving objects in the environment. In the indoor environment, besides the human target, there are also other objects which may generate dynamic reflected path signals, such as pets and electric fan. The moving

object with a constant speed, such as the electric fan, will incur a dynamic reflected path signal with a fixed path change speed. On the other hand, the path change speed induced by the human target changes with time. Thus, the interference from a constant-speed moving object can be mitigated by removing the component with a fixed path change speed. However, the dynamic reflected path induced by a moving pet is similar to that caused by a moving human. Distinguishing dynamic reflected path signals incurred by a moving pet and a non-still human target is still challenging and it will be an important part of our future work.

Supporting more Wi-Fi based device-free applications. Besides activity recognition, the vitality information captured by our platform could also be used for other Wi-Fi based device-free applications. For example, Wi-Fi signal can also be used for human respiration monitoring [22, 35]. The subtle movement of human chest incurred by the respiration will also make the target reflected path length change. We believe the path change speed spectrum estimated by our platform may be used for fine-grained respiration sensing in the future. Moreover, if we want to monitor the sleep respiration, our platform could detect whether the target is sleeping in the bedroom as a trigger of human respiration detection. In the future, we plan to develop more sensing applications on top of our platform to provide richer sensing information about the target without any extra deployment.

9 RELATED WORK

Our work is broadly related to research in the areas of indoor human motion detection and localization. Over the years, many different indoor human sensing technologies have been developed including camera [5, 7, 26], sound [24, 41], radio frequency [2, 27], inertial sensors [4, 45], infrared [11, 14], visible light [13, 17], ambient sensors [6, 12, 23, 25, 30] and passive infrared (PIR) motion sensor [34, 53]. Camera-based technologies require good lighting conditions and raise privacy concerns. Sound-based technologies are vulnerable to acoustic noise and the coverage area is limited. Inertial sensor-based methods require the target to carry or wear a device for sensing. Dedicated infrastructure is needed for infrared based system while visible light only works in LoS scenarios. Ambient sensor based solutions usually require a dense deployment which incurs high costs for installation and maintenance. Compared with PIR motion sensor, our platform not only could detect when the target is non-still and which area the target is staying in, but could also capture the movement speed characteristics of the human target, which enables more fine-grained sensing applications. Moreover, Wi-Fi signal has already been used for respiration detection [22, 35], gesture recognition [18, 18], etc. We can further build these sensing applications on top of our platform to provide more sensing information without any extra device. Therefore, in this work, we focus on device-free human vitality sensing using the commodity Wi-Fi devices that already exist at home. We discuss the most related research works here.

9.1 Wi-Fi Based Indoor Localization

In the last few years, Wi-Fi based indoor localization draws a lot of attention from both academia and industry. Many Wi-Fi based localization systems have been developed, including both device-based and device-free solutions. For device-based solutions, earlier works mostly utilize RSSI (Received Signal Strength Indicator) and achieve meter-level accuracies [3, 51]. Recently, adopting CSI and antenna array, decimeter-level localization has been achieved [16, 49]. However, these solutions require the target to hold a device for localization.

Since Youssef et al. [52] introduced the concept of device-free localization in 2007, many device-free localization systems have been proposed. Nuzzer [33] and Ichnaea [31] use the RSSI signature as a fingerprint for localization. Pilot [48] and MonoPHY [1] systems employ the finer CSI information as the fingerprint to improve the accuracy. LiFS [38] utilizes fresnel model to improve the accuracy of localization when the target is on the LoS path of the transceivers. Most of these works need labor intensive offline training to build the fingerprint database, which needs to be updated when the environment changes. Moreover, these works can not detect whether the human is non-still. The most relevant work to WiVit is E-eyes [42]. E-eyes utilizes the amplitude pattern of CSI to build a

fingerprint map to identify the target's moving trajectory and detect whether the target is non-still. However, the E-eyes system also requires human efforts in offline training and calibration for different environments.

In the last few years, model-based device-free localization solutions have also been developed. The Dynamic-MUSIC method [20] and IndoTrack system [21] employ the Angle-of-Arrival (AoA) information to locate and track the target. However, due to the random phase offset between two RF ports, AoA-based solutions require careful phase calibration during startup [49]. MFDL [36] employs fresnel model to locate the target during walking but also requires offline training to calibrate the position of each fresnel zone in the multipath environment. Due to lack of movement direction information, WiDar [29] tracks a target based on the amplitude of Doppler frequency shift but only considers a single area.

Compared with these works, our WiVit does not require the target to hold any device and can identify when the target is non-still and in which area the target is staying in a multi-room environment without any human effort in offline training or calibration.

9.2 Wi-Fi Based Human Motion Detection

Wi-Fi signal has also been employed to detect human motion. Earlier solutions detect the human motion based on the variance of RSSI [15, 52]. In the last few years, fine-grained CSI has been used for human motion detection. CARM [40] employs the variance of CSI amplitude to detect motion for activity recognition. FIMD [47] utilizes the correlation of CSI amplitude over time for motion detection. PADS [28] combines both phase and amplitude information of CSI to improve motion detection accuracy. DeMan [46] not only detects the human motion but also detects the existence of a human even if the target is still based on the human respiration sensing. MoSense [9] and AR-Alarm [19] utilize the variance of phase difference for human motion detection. RT-Fall [37] also utilizes the signal phase difference to detect when the human is still for fall detection. All these works require offline training and calibration to obtain a threshold value to distinguish the still and non-still status of the target and this threshold varies in different environments. On the other hand, our WiVit platform detects the target's non-still status by capturing the constant CSI phase change of the target reflected path signal incurred by target movements. This feature is only related to the target reflected path signal and independent of the environmental noise and multipath. Thus, our platform is robust against different environments without any human intervention.

10 CONCLUSION

In this paper, we propose a device-free human vitality sensing platform WiVit hosted on cheap commodity Wi-Fi devices. The WiVit platform can accurately capture when the target is non-still and in which area the target is staying without offline training or calibration, moving one step further towards real life adoptions. For human activity detection, the precision is 98% and the FNR is as low as 1%. Meanwhile, based on the relationship between human target's position and the reflected path change speed at each receiver, WiVit could identify which area the human target is staying in at an accuracy close to 100%. We also employ two use cases to show the potential applications of our platform. We deploy an activity recognition system on top of our WiVit platform with an average accuracy of 94.2%. We further conduct a feasibility study to show that the captured human vitality statistics could accurately reflect the the unique characteristic of each daily activity and thus can be used for long-term daily life monitoring. We believe our platform can provide valuable datasets to infer the high-level semantics of one's different daily life facets such as living habits, physical conditions and even emotions.

ACKNOWLEDGMENTS

The work is supported by the National Key Research and Development Plan under Grant No. 2016YFB1001200 and Peking University Information Technology Institute (Tianjin Binhai).

REFERENCES

[1] Heba Abdel-Nasser, Reham Samir, Ibrahim Sabek, and Moustafa Youssef. 2013. MonoPHY: Mono-stream-based device-free WLAN localization via physical layer information. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)*. 4546–4551.

[2] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C. Miller. 2014. 3D Tracking via Body Radio Reflections. In *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI '14)*. USENIX Association, Seattle, WA, 317–329.

[3] Paramvir Bahl and Venkata N. Padmanabhan. 2000. RADAR: an in-building RF-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, Vol. 2. IEEE, 775–784 vol.2.

[4] Agata Brajdic and Robert Harle. 2012. Scalable Indoor Pedestrian Localization Using Inertial Sensing and Parallel Particle Filters. In *Proceedings of 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN '12)*. IEEE, 1–10.

[5] Q. Cai and J. K. Aggarwal. 1998. Automatic Tracking of Human Motion in Indoor Scenes across Multiple Synchronized Video Streams. In *Proceedings of the Sixth International Conference on Computer Vision (ICCV '98)*. IEEE, 356–362.

[6] Liming Chen and Hui Nugent, Chris D.and Wang. 2012. A Knowledge-Driven Approach to Activity Recognition in Smart Homes. *IEEE Transactions on Knowledge and Data Engineering* 24, 6 (June 2012), 961–974.

[7] Mohamed Eldib, Francis Deboeverie, Wilfried Philips, and Hamid Aghajan. 2016. Behavior Analysis for Elderly Care Using A Network of Low-Resolution Visual Sensors. *J. Electron. Imaging* 25, 4 (Mar 2016), 17.

[8] Tao Gu, Liang Wang, Zhanqing Wu, Xianping Tao, and Jian Lu. 2011. A Pattern Mining Approach to Sensor-Based Human Activity Recognition. *IEEE Transactions on Knowledge and Data Engineering* 23, 9 (Sept 2011), 1359–1372.

[9] Yu Gu, Jinhai Zhan, Yusheng Ji, Jie Li, Fuji Ren, and Shangbing Gao. 2017. MoSense: An RF-Based Motion Detection System via Off-the-Shelf WiFi Devices. *IEEE Internet of Things Journal* 4, 6 (Dec 2017), 2326–2341.

[10] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool Release: Gathering 802.11N Traces with Channel State Information. *SIGCOMM Comput. Commun. Rev.* 41, 1 (Jan. 2011), 53–53.

[11] Ju Han and Bir Bhanu. 2005. Human Activity Recognition in Thermal Infrared Imagery. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05) - Workshops - Volume 03 (CVPR '05)*. IEEE, Washington, DC, USA, 17–17.

[12] Derek Hao Hu, Sinno Jialin Pan, Vincent Wenchen Zheng, Nathan Nan Liu, and Qiang Yang. 2008. Real World Activity Recognition with Multiple Goals. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*. ACM, New York, NY, USA, 30–39.

[13] Pan Hu, Liqun Li, Chunyi Peng, Guobin Shen, and Feng Zhao. 2013. Pharos: Enable Physical Analytics Through Visible Light Based Indoor Localization. In *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks (HotNets-XII)*. ACM, New York, NY, USA, Article 5, 7 pages.

[14] Jung-Yoon Kim, Na Liu, Hwee-Xian Tan, and Chao-Hsien Chu. 2017. Unobtrusive Monitoring to Detect Depression for Elderly With Chronic Illnesses. *IEEE Sensors Journal* 17, 17 (Sept 2017), 5694–5704.

[15] Ahmed E. Kosba, Ahmed Saeed, and Moustafa Youssef. 2012. RASID: A robust WLAN device-free passive motion detection system. In *2012 IEEE International Conference on Pervasive Computing and Communications*. 180–189.

[16] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using Wi-Fi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM '15)*. ACM, New York, NY, USA, 269–282.

[17] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor Positioning with Mobile Phones and Visible Light. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom '14)*. ACM, New York, NY, USA, 447–458.

[18] 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 (UbiComp '16)*. ACM, New York, NY, USA, 250–261.

[19] Shengjie Li, Xiang Li, Kai Niu, Hao Wang, Yue Zhang, and Daqing Zhang. 2017. AR-Alarm: An Adaptive and Robust Intrusion Detection System Leveraging CSI from Commodity Wi-Fi. In *International Conference on Smart Homes and Health Telematics (ICOST '17)*. Springer International Publishing, 211–223.

[20] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. 2016. Dynamic-MUSIC: Accurate Device-free Indoor Localization. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 196–207.

[21] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. 2017. IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 72 (Sept. 2017), 22 pages.

[22] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2016. Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices. *IEEE Transactions on Mobile Computing* 15, 10 (Oct 2016), 2466–2479.

[23] Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. 2007. A Long-Term Evaluation of Sensing Modalities for Activity Recognition. In *UbiComp 2007: Ubiquitous Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 483–500.
[24] Wenguang Mao, Jian He, and Lili Qiu. 2016. CAT: High-precision Acoustic Motion Tracking. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 69–81.
[25] Elizabeth D. Mynatt, Jim Rowan, Sarah Craighill, and Annie Jacobs. 2001. Digital Family Portraits: Supporting Peace of Mind for Extended Family Members. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 333–340.
[26] Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, and Tatsuya Harada. 2016. Recognizing Activities of Daily Living With a Wrist-Mounted Camera. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR '16)*. 3103–3111.
[27] 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 and Networking (MobiCom '13)*. ACM, New York, NY, USA, 27–38.
[28] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Zimu Zhou. 2014. PADS: Passive detection of moving targets with dynamic speed using PHY layer information. In *2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*. 1–8.
[29] Kun Qian, Chenshu Wu, Zheng Yang, Chaofan Yang, and Yunhao Liu. 2016. Decimeter Level Passive Tracking with Wi-Fi. In *Proceedings of the 3rd Workshop on Hot Topics in Wireless (HotWireless '16)*. ACM, New York, NY, USA, 44–48.
[30] Daniele Riboni, Timo Sztyler, Gabriele Civitarese, and Heiner Stuckenschmidt. 2016. Unsupervised Recognition of Interleaved Activities of Daily Living Through Ontological and Probabilistic Reasoning. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 1–12.
[31] Ahmed Saeed, Ahmed E. Kosba, and Moustafa Youssef. 2014. Ichnaea: A Low-Overhead Robust WLAN Device-Free Passive Localization System. *IEEE Journal of Selected Topics in Signal Processing* 8, 1 (Feb 2014), 5–15.
[32] R. Schmidt. 1986. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Transactions on Antennas and Propagation* 34, 3 (Mar 1986), 276–280.
[33] Moustafa Seifeldin, Ahmed Saeed, Ahmed E. Kosba, Amr El-Keyi, and Moustafa Youssef. 2013. Nuzzer: A Large-Scale Device-Free Passive Localization System for Wireless Environments. *IEEE Trans. Mobile Computing* 12, 7 (July 2013), 1321–1334.
[34] Byunghun Song, Haksoo Choi, and Hyung Su Lee. 2008. Surveillance Tracking System Using Passive Infrared Motion Sensors in Wireless Sensor Network. In *2008 International Conference on Information Networking*. 1–5.
[35] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. 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 (UbiComp '16)*. ACM, New York, NY, USA, 25–36.
[36] Hao Wang, Daqing Zhang, Kai Niu, Qin Lv, Yuanhuai Liu, Dan Wu, Ruiyang Gao, and Bing Xie. 2017. MFDL: A Multicarrier Fresnel Penetration Model based Device-Free Localization System leveraging Commodity Wi-Fi Cards. *CoRR* abs/1707.07514 (2017). arXiv:1707.07514
[37] Hao Wang, Daqing Zhang, Yasha Wang, Junyi Ma, Yuxiang Wang, and Shengjie Li. 2017. RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi Devices. *IEEE Transactions on Mobile Computing* 16, 2 (Feb 2017), 511–526.
[38] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. 2016. LiFS: Low Human-effort, Device-free Localization with Fine-grained Subcarrier Information. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 243–256.
[39] Wei Wang, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and Modeling of Wi-Fi Signal Based Human Activity Recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 65–76.
[40] Wei Wang, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2017. Device-Free Human Activity Recognition Using Commercial WiFi Devices. *IEEE Journal on Selected Areas in Communications* 35, 5 (May 2017), 1118–1131.
[41] Wei Wang, Alex X. Liu, and Ke Sun. 2016. Device-free Gesture Tracking Using Acoustic Signals. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 82–94.
[42] 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 the 20th Annual International Conference on Mobile Computing and Networking (MobiCom '14)*. ACM, New York, NY, USA, 617–628.
[43] Lloyd R. Welch. 2003. Hidden Markov Models and the Baum-Welch Algorithm. 53, 4 (Dec 2003), 10–13.
[44] Wi-Fi Alliance® 6 for '16 Wi-Fi® predictions 2016. Retrieved January 31, 2018 from https://www.wi-fi.org/beacon/wi-fi-alliance/wi-fi-alliance-6-for-16-wi-fi-predictions
[45] Oliver Woodman and Robert Harle. 2009. RF-Based Initialization for Inertial Pedestrian Tracking. In *Proceedings of the 7th International Conference on Pervasive Computing (Pervasive '09)*. Springer-Verlag, Berlin, Heidelberg, 238–255.
[46] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. 2015. Non-Invasive Detection of Moving and Stationary Human With WiFi. *IEEE Journal on Selected Areas in Communications* 33, 11 (Nov 2015), 2329–2342.

[47] Jiang Xiao, Kaishun Wu, Youwen Yi, Lu Wang, and Lionel M. Ni. 2012. FIMD: Fine-grained Device-free Motion Detection. In *2012 IEEE 18th International Conference on Parallel and Distributed Systems*. 229–235.

[48] Jiang Xiao, Kaishun Wu, Youwen Yi, Lu Wang, and Lionel M. Ni. 2013. Pilot: Passive Device-Free Indoor Localization Using Channel State Information. In *2013 IEEE 33rd International Conference on Distributed Computing Systems*. 236–245.

[49] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-grained Indoor Location System. In *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI '13)*. USENIX Association, Berkeley, CA, USA, 71–84.

[50] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. 2015. ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 537–549.

[51] Moustafa Youssef and Ashok Agrawala. 2005. The Horus WLAN Location Determination System. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys '05)*. ACM, New York, NY, USA, 205–218.

[52] Moustafa Youssef, Matthew Mah, and Ashok Agrawala. 2007. Challenges: Device-free Passive Localization for Wireless Environments. In *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '07)*. ACM, New York, NY, USA, 222–229.

[53] Zhaoyuan Yu, Linwang Yuan, Wen Luo, Linyao Feng, and Guonian Lv. 2016. Spatio-Temporal Constrained Human Trajectory Generation from the PIR Motion Detector Sensor Network Data: A Geometric Algebra Approach. *Sensors* 16, 1 (2016).

[54] Fusang Zhang, Daqing Zhang, Jie Xiong, Hao Wang, Kai Niu, Beihong Jin, and Yuxiang Wang. 2018. From Fresnel Diffraction Model to Fine-grained Human Respiration Sensing with Commodity Wi-Fi Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 53 (March 2018), 23 pages.

Received February 2018; revised May 2018; accepted August 2018