

# Modeling RFID Signal Reflection for Contact-free Activity Recognition

YANWEN WANG, The Hong Kong Polytechnic University

YUANQING ZHENG\*, The Hong Kong Polytechnic University

Wireless sensing techniques for tracking human activities have been vigorously developed in recent years. Yet current RFID based human activity recognition techniques need either direct contact to human body (e.g., attaching RFIDs to users) or specialized hardware (e.g., software defined radios, antenna array). How to wirelessly track human activities using commodity RFID systems without attaching tags to users (i.e., a contact-free scenario) still faces lots of technical challenges. In this paper, we quantify the correlation between RF phase values and human activities by *modeling intrinsic characteristics of signal reflection in contact-free scenarios*. Based on the signal reflection model, we introduce TACT that can recognize human activities using commodity RFIDs without attaching any RFID tags to users. TACT first reliably detects the presence of human activities and segments phase values. Then, candidate phase segments are classified according to their coarse-grained features (e.g., moving speed, moving distance, activity duration) as well as their fine-grained feature of phase waveform. We deploy and leverage multiple tags to increase the coverage and enhance the robustness of the system. We implement TACT with commodity RFID systems. We invite 12 participants to evaluate our system in various scenarios. The experiment results show that TACT can recognize eight types of human activities with 93.5% precision under different and challenging experiment settings.

CCS Concepts: • **Networks** → **Network mobility**; • **Human-centered computing** → **Ubiquitous and mobile computing**; • **Computer systems organization** → Sensor networks;

Additional Key Words and Phrases: Activity recognition, RFID systems, contact-free, backscatter communication

## ACM Reference Format:

Yanwen Wang and Yuanqing Zheng. 2018. Modeling RFID Signal Reflection for Contact-free Activity Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 193 (December 2018), 22 pages. <https://doi.org/10.1145/3287071>

## 1 INTRODUCTION

Recent years have witnessed a surge of wireless and mobile devices and their applications. Human activity recognition enables user interaction with smart environment, monitors users' daily activities, and infers their health conditions. For instance, by monitoring activities of elderly people, one may track their well-being and behavioral changes, and provide timely help in case of accidents or injuries [3]. Recent works try to leverage various commodity devices available on the market to build activity recognition systems so as to enable such applications.

\*This is the corresponding author

Authors' addresses: Yanwen Wang, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, yanwen.wang@connect.polyu.hk; Yuanqing Zheng, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, csyqzheng@comp.polyu.edu.hk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

2474-9567/2018/12-ART193 \$15.00

<https://doi.org/10.1145/3287071>

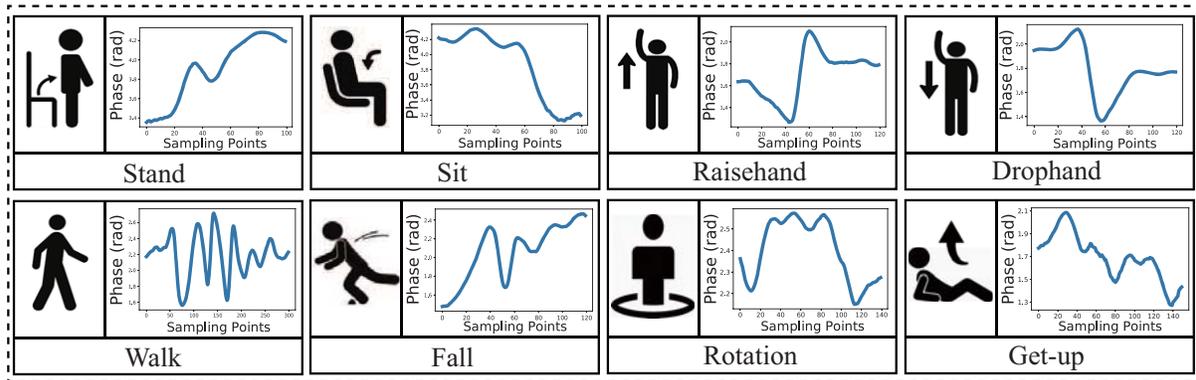


Fig. 1. Eight types of human activities.

Existing activity recognition methods primarily focus on cameras, motion sensors, or other specialized sensors. These traditional methods need either Line-of-Sight (LOS) path (e.g., camera-based techniques) or direct body contact (e.g., sensor-based techniques), which might raise privacy concerns and cause inconvenience to users. Recent wireless sensing technologies explore the potential of enabling activity recognition using wireless signals. The intuition is that human body may influence wireless signals in the environment and cause changes in signal strength and phase. By identifying unique patterns induced by different activities, one may infer different types of human activities. Capturing such changes using wireless devices (e.g., RFID, WiFi) however is challenging in practice. Deploying more wireless devices can possibly enrich wireless links in the environment but inevitably incurs extra deployment cost. Moreover, many wireless based methods [1, 2, 22, 23] need specialized radios (e.g., USRP [22, 23], FMCW radio [1, 2]), which are not readily available on the market.

Commodity RFID systems are widely used in object identification and tracking [10, 16, 17, 20, 24, 27, 39, 40] and provide opportunity to infer human activities from RFID-labeled objects [5, 28, 33]. Unlike traditional radio based methods, the RFID based methods can enhance the coverage by deploying multiple tags at very low marginal cost of passive RFID tags. Daily objects (e.g., door) can be labeled with RFID tags which allow Tagyro [33] to infer activities (e.g., door opening) from the orientation changes of the labeled objects. TagBreathe [9] attaches RFID tags to users and measures the tiny movement of human chest to monitor respiration. Note that such systems need to attach RFID tags to specific objects or users to track their movement and infer different types of activities. Yet, many activities (e.g., walking, push up, body rotation, etc.) do not directly interact with RFID-labeled objects or cause movement of those objects in the environment. Empirical results have shown that RFID signals can be influenced by nearby human activities (e.g., blockage of line-of-sight path) even if the objects are not moved. Recent works adopt deep learning based methods to automatically analyze the signal changes due to human activities and extract meaningful features from the RFID signals [6, 14]. Such works need a large number of training data to train deep neural networks [14] and are susceptible to environment changes. For instance, TagFree [6] uses customized antenna arrays to measure Angle of Arrival (AoA) of the multipath signals and processes the AoA measurements with neural networks. In case of environment changes, the neural networks need to be updated and re-trained with new data. Thus, the data collection and training overhead is high in practice.

Note that unlike previous backscatter propagation models [9, 39] which quantify the relative *movement between a tag and a reader*, in contact-free scenarios *both devices may keep motionless and untouched* in activity recognition process. To the best of our knowledge, we still lack a signal reflection model for RFID based human activity recognition in the contact-free scenarios. As such, existing works try to adopt neural network models to

automatically extract features. The machine extracted features however may not always be able to differentiate various human activities, especially when the training dataset is small.

To fill this gap, we propose a contact-free reflection model for activity recognition based on commodity RFID systems and extract meaningful features for human activities based on the model.

(1)*Contact-free reflection model for activity recognition based on commodity RFID systems*: We build a contact-free reflection model by conducting in-depth theoretical analysis of RFID signals in contact-free scenarios. We validate the model with a series of field experiments on commodity RFID systems. The contact-free reflection model allows us to quantitatively correlate the measured RFID phase values with different human activities.

(2)*Feature extraction for human activities based on the contact-free reflection model*: We leverage the contact-free reflection model and design intrinsic features for contact-free activity recognition. In particular, we extract both coarse-grained and fine-grained features to capture subtle differences of human activities in practice. Fig. 1 depicts the eight different types of activities and the typical phase waveforms measured during the corresponding activities. We quantify the similarity of phase waveforms using the Dynamic Time Warping (DTW) algorithm to compensate for the temporal differences. Based on the contact-free reflection model, we also extract other meaningful features such as moving speed and moving distance which are meaningful to differentiate various human activities. We also conduct optimization and improve system robustness based on the model.

We implement TACT with commodity RFID systems and conduct extensive evaluation in various experiment settings. In our implementation, we adopt a simple yet effective segmentation algorithm to identify activity corresponded phase segments. We deploy multiple tags and combine their readings to improve the system coverage and enhance the system robustness in case of blockage of line-of-sight path and signal attenuation. We evaluate the performance of extracted coarse-grained and fine-grained features by training and comparing multiple classifiers (e.g., neural networks, decision tree, support vector machine, random forest, etc) based on the Scikit-learn library [21]. We invited 12 different volunteers to perform 8 activities multiple times in 3 different rooms. We conducted the experiments at different times across 3 months. We collected 640 human activities for training and 1280 for testing. The evaluation results show that TACT achieves high detection accuracy of 93.5% even with a relatively small number of training data.

## 2 MODELING RFID REFLECTION FOR DEVICE-FREE ACTIVITY RECOGNITION

### 2.1 RFID System

An RFID system consists of two parts, a high-power reader and a batteryless tag. The reader transmits a continuous wave to energize the tag, and the tag responds by backscattering the continuous wave. The modulation is based on on-off keying, where the tag either reflects or absorbs the incoming continuous wave. The reader decodes the tag's response and measures physical layer information such as signal phase and signal strength. Suppose the reader transmits the continuous wave and meanwhile receives reflected signal of the tag. The received signals at the reader can be simply classified as either *with-tag-reflection* or *without-tag-reflection*. Since without-tag-reflection signals are not modulated by any tag, they can be regarded as delayed copies of continuous waves, which will be filtered out by self-interference cancellation circuits of commodity RFID readers [40]. Such self-interference cancellation modules are essential to the decoding performance of commodity RFID systems. As such, we can safely ignore the without-tag-reflection in the system model. The reader can demodulate the with-tag-reflection signal and provide the corresponding low-level information (e.g., signal strength, phase).

### 2.2 Understanding the Reflection of RFID Signal

To understand the signal reflection from a moving object and build a reflection model for contact-free scenarios, we conducted preliminary experiments using the Impinj R420 reader and the commodity passive tag. As shown

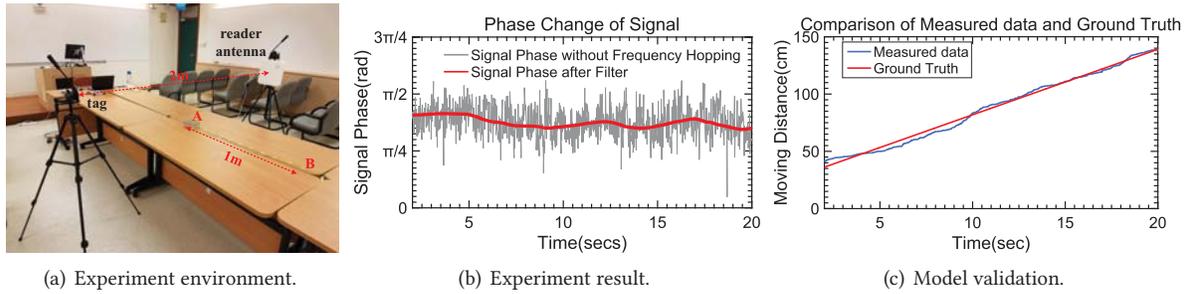


Fig. 2. Metal Object Moving along the Perpendicular bisector between Reader and Tag.

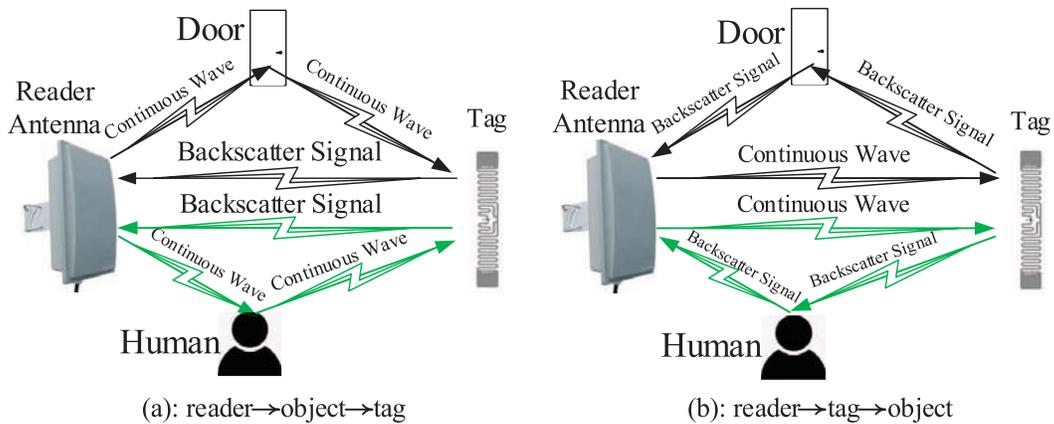


Fig. 3. The signals that are reflected two times.

in Fig.2(a), the distance between the reader’s antenna<sup>1</sup> and the tag was around 2m. Both the tag and the reader’s antenna remained stationary throughout the experiments. We used a metal reflector to reflect wireless signals in the environment. The metal reflector moved at a constant speed of 5cm/s along the perpendicular bisector between the tag and the antenna and the moving distance was around 1m. Note that if the signal propagation distance increases one wavelength, the signal phase changes  $2\pi$  as described in [39]. As the metal reflector moved at a constant speed, the propagation distance of signal reflected from the metal reflector increased linearly. Thus we may expect to observe a linear and discontinuous phase change ranging from 0 to  $2\pi$  if a tag were attached to the box.

In the contact-free scenarios, however, as the red line plotted in Fig.2(b), the received phase values after applying a median filter were different from what we expected. The phase values continuously changed and exhibited a non-linear pattern. The range of phase values was only around 0.2, which was much smaller than  $2\pi$ . The experiment demonstrates that it is indeed possible to observe some changes in phase values when a reflector moves around the RFID system even though both the tag and the reader remain stationary. Yet, it still remains elusive why and how such a waveform was generated.

<sup>1</sup>In the following, we use reader’s antenna and reader interchangeably.

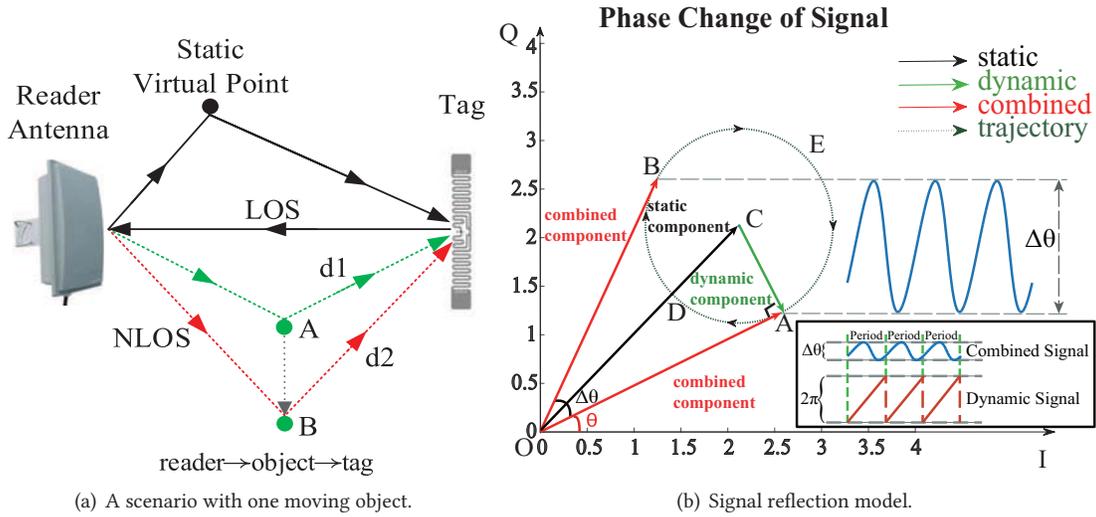


Fig. 4. The change of superimposed signal phase when an object moves in the environment.

**2.2.1 Signal Phase in Static Environment.** To answer this question, we start by analyzing RF signal phase change in the static environment. Note that the without-tag-reflection signals can be regarded as self-interference filtered out by the self-interference cancellation module of commodity RFID systems. On the other hand, the with-tag-reflection signals are modulated by tags and will be extracted by the reader. In our analysis, we focus on the  $k$ -reflection signals ( $k \leq 2$ ), i.e., signals reflected at most 2 times before arriving at the reader. This is because due to the increase of propagation paths and signal attenuation, the wireless signals become orders of magnitude weaker after multiple reflections in the environment. Thus, we regard other reflection signals (e.g., 4-reflection signals) as noise and interference in our model.

Fig.3 depicts the with-tag-reflection signals that are reflected two times by a tag and an object (e.g., a door, a user). Each object introduces two possible propagation paths in the space as illustrated in Fig.3. Fig.3(a) depicts reader → object → tag → reader and Fig.3(b) depicts reader → tag → object → reader, respectively. Since these signals experience two reflections, we define these signals as *2-reflection signals*. The phase of each signal depends on the distance of signal propagation path in the space, as follows:

$$\theta = (2\pi \times \frac{d_{propagation}}{\lambda} + \hat{\theta}) \bmod 2\pi \quad (1)$$

where  $d_{propagation}$  represents the signal propagation distance.  $\lambda$  is the wavelength of the signal and  $\hat{\theta}$  is the constant phase rotation introduced by the hardware of reader and tag [39]. Note that although one object (e.g., human) introduces two propagation paths (e.g., reader-human-tag-reader, reader-tag-human-reader), these two propagation paths are *reciprocal* in the space with the same propagation distance. As a result, the superimposed signal will be the sum of the two signals with the same phase as the 2-reflection signals. Without loss of generality, in our following analysis, we only consider one path for each object (e.g., reader-object-tag-reader).

In an environment without any moving object, the received signals are reflected from static objects (e.g., walls, desks). Since the propagation distance from a static object remains constant, the phase of the corresponding signal

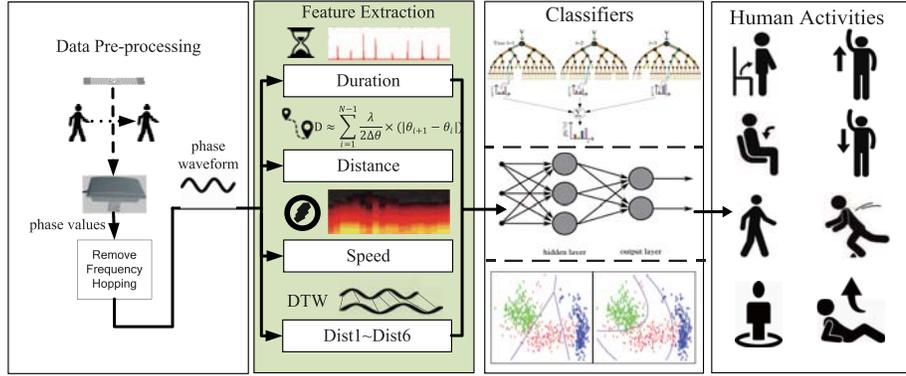


Fig. 5. System architecture of TACT.

does not change. Thus, if the environment remains static, the phase of superimposed signal will be constant. In our model, we use a *virtual static reflection point* to represent the static environment, as shown in Fig.4(a).

**2.2.2 Signal Phase Introduced by Moving Objects.** Suppose a moving object moves from A to B as in Fig.4(a). The distance of the 2-reflection signal will change (e.g., from reader→A→tag→reader to reader→B→tag→reader as depicted in Fig.4(a)), which results in the phase change of received signals.

Fig.4(b) illustrates how the phase of superimposed signal changes when an object moves in a relatively static environment. Specifically, two components jointly determine the phase of finally received superimposed signal. The first component is the static component, which represents the *virtual static reflection*. The phase of the static component will remain constant. The second component is the dynamic component, which represents the reflected signals from the moving object. As illustrated in Fig.4(b),  $\vec{OC}$  represents the static component, while  $\vec{CA}$  represents the dynamic component. As the object moves, the dynamic component  $\vec{CA}$  may rotate  $2\pi$ , while the static component  $\vec{OC}$  remains constant. As a result, we may observe the smallest phase and the largest phase of combined signal at the two tangent points A and B on the circle, respectively. The phase difference  $\angle AOB$  is the phase change range.

In Fig.4(b), we have three key observations. First, the phase of combined signal will change over a range of  $[\theta, \theta + \Delta\theta]$ ,  $\Delta\theta \leq 2\pi$ . Second, suppose the dynamic component rotates clockwise for  $2\pi$  at a constant angular speed. Since  $\widehat{ADB}$  is shorter than  $\widehat{BEA}$ , the phase waveform may not exhibit as a standard sinusoid form. Third, if the object keeps moving and the dynamic component keeps rotating, we may observe periodic phase values of the combined signal. The combined signal and dynamic signal share the same period since both of them rotate  $2\pi$  for one period. Due to the same period, we may quantify the correlation by directly using the combined signal phase values rather than dynamic signal, which is very hard to acquire.

**2.2.3 Model Validation.** To empirically validate our model, we measure the total moving distance of the object from the phase waveform in Fig.2(b) by measuring the phase changes between 2 consecutive phase readings throughout the moving period. Note that when the propagation distance changes one wavelength, the phase will change  $2\Delta\theta$  in our model. Fig.2(c) shows the measured results according to our model. The measured distance almost overlaps with the ground truth. The speed of the object can be calibrated by the slope of the measured data, which is  $5\text{cm/sec}$ , the same as our experiment configuration.

### 3 SYSTEM DESIGN

Based on the contact-free backscatter model as illustrated in Fig.4(b), we design and implement an activity recognition system named TACT in this section. Fig.5 depicts the architecture of TACT. TACT performs data preprocessing by mitigating frequency hopping effects on the phase readings collected from the tag. Both coarse-grained and fine-grained features are extracted from the preprocessed data. TACT combines these coarse-grained and fine-grained features to recognize different human activities. We describe the key components of TACT in detail.

#### 3.1 Data Pre-processing

In data pre-processing, the first step is to eliminate the discontinuity of phase readings caused by frequency hopping in commodity RFID systems. The carrier frequency hops among multiple sub-channels to avoid co-channel interference periodically, which results in discontinuity of phase readings. To mitigate the impact of frequency hopping, we map different hopping frequencies to a single fixed frequency as described in [33]:

$$\theta_m = \left[ \frac{f_n}{f_k} \times (\theta_{d,k} - \theta_{d_0,k}) + \theta_{d_0,n} \right] \bmod 2\pi \quad (2)$$

where  $f_n$  is the fixed frequency of a specific channel (e.g.,  $f_n = 920.25\text{MHz}$  in our experiment),  $f_k$  is the current carrier frequency of the signal derived from channel  $k$ .  $\theta_{d_0,n}$  and  $\theta_{d_0,k}$  are the initial phase values of channel  $n$  and channel  $k$  collected in static environment with an arbitrary constant distance  $d_0$  between reader and tag.  $\theta_{d,k}$  is the phase of the current signal from channel  $k$ . All the parameters in Eq.(2) can be directly measured and the calibration overhead is negligible. With this method, all the phase values with different carrier frequencies are normalized to the phases with a fixed carrier frequency, as if the phases were measured without frequency hopping.

Due to the environmental noise, even though the channel frequency hopping effect is mitigated, it is still very hard to observe clear pattern induced by the moving object. TACT uses a median filter to remove the outliers in the phase readings. A median filter replaces each phase reading with the median value within a range of neighboring samples, which can appropriately filter out the spikes in the data sequence caused by sudden noise. Empirically, TACT sets the window size of moving median filter to 1 second.

To identify phase segments corresponding to human activities, we design a simple yet effective segmentation algorithm. Our intuition is that the variance of phase readings can serve as a good indicator for activity segmentation. During a period of 100s, a user walks for a few steps, stops walking for a while, and repeats 8 times. In Fig.6(a), we see that the phase readings fluctuate when the user walks, and remain relatively stable when the user stops. A threshold for the phase readings cannot reliably segment the activities due to noise and interference in the environment. In contrast, we consistently observe large increases in variances during the 8 walking periods as shown in Fig.6(b). Therefore, we compute the variances for every window of 8 phase readings. If the variances exceed a threshold for 3 consecutive windows, the segmentation algorithm starts to record the phase readings for possible activities and stops when the variances drop below the threshold for 3 consecutive windows. For the threshold, generally a larger threshold would filter out noises while missing weak phase changes. In practice, we use an empirical threshold of 3 standard deviation (i.e., 99.7% confidence level) to detect and segment activities [38].

#### 3.2 Feature Extraction

In this section, based on the reflection model, we extract key features from phase values that differentiate human activities.

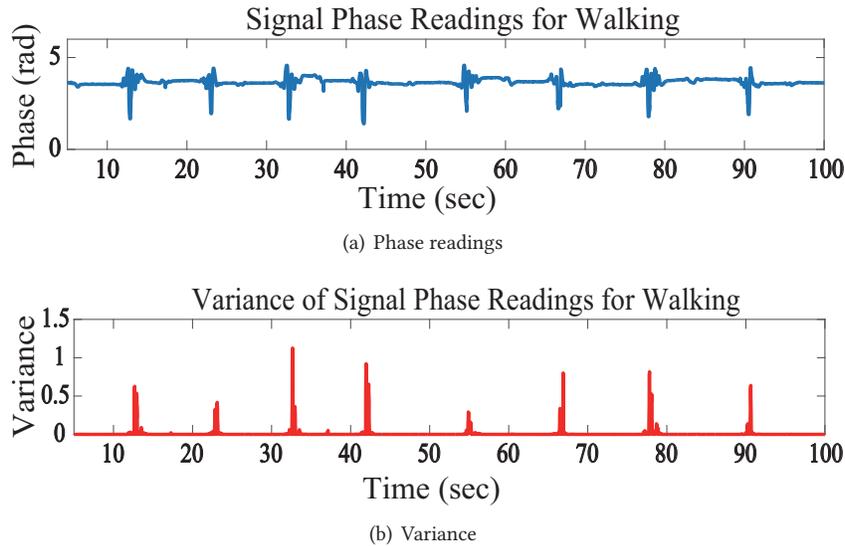


Fig. 6. Signal phase readings and corresponding variance.

**3.2.1 Phase Waveform.** Our empirical study shows that the signal phases exhibit similar waveforms when the same activities are performed even by different persons, while different activities lead to different phase waveforms. In our experiment, we invited three volunteers to perform 3 different activities including raising-hand, walking and falling-down within the sensing range of the RFID system in a large room at different times. The arrangement of furniture was different to investigate the influence of environment changes in the same room. Each activity was performed 30 times by each user. We randomly choose 3 phase waveforms for each activity performed by the 3 users as illustrated in Fig.7. We observe that different activities exhibit different waveforms. In contrast, the same activities exhibit similar waveforms, even when the same activities were performed by different users. However, the phase waveforms of the same activities may be *misaligned* due to the differences in time duration and speed of the activities. The reason for such *misalignment* lies on two-fold: (1) Due to channel contentions among tags, the collected data and the reference data may not be synchronized since RFID systems adopt Slotted ALOHA protocol. If the collected data is from a tag with a lower reading rate, the measured phase waveform would appear to be squeezed compared with a reference waveform, while a tag with a higher sampling rate would stretch the measured phase waveform compared with the same reference waveform. (2) Although people perform the same activities, different people have different torso speeds and distances due to different genders, ages, and statures. Even when the same person performs the same activities, we may observe slightly misaligned phase waveforms.

Dynamic Time Warping (DTW) algorithm is able to handle such misalignment with restriction and measure the similarity or distance between two phase waveforms [27]. Therefore, we adopt DTW to calculate the distance between an unknown phase waveform and reference waveforms. Before the DTW measurement, we first remove the DC offset of phase waveforms such that the calculated distance only quantifies the shape discrepancy between the waveforms. It is worth noting that the environmental changes caused by different layouts are eliminated by removing the DC offset such that TACT is resilient to the environmental changes. We further normalize the

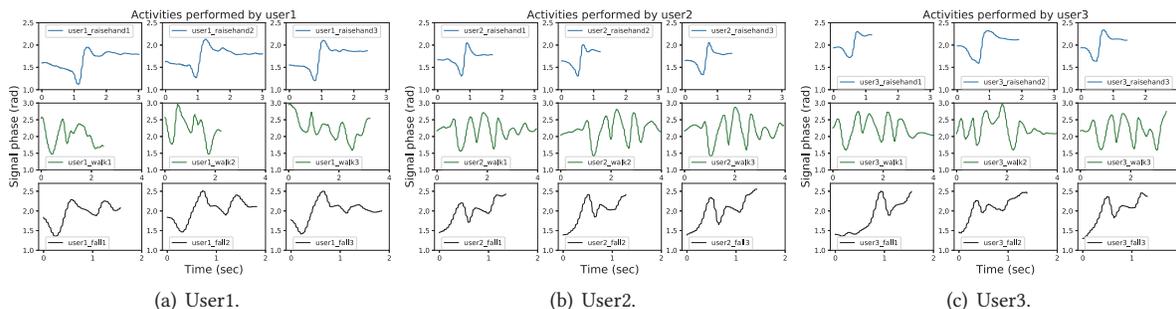


Fig. 7. Signal phase waveforms of different activities performed by different volunteers.

distance value with the number of matching pairs in DTW calculation. The activity type is identified with the smallest distance to template phase waveform of each activity, which indicates the highest similarity.

DTW incurs relatively high computation overhead of  $O(n \times m)$ , where  $n$  and  $m$  are the numbers of data in the data sequences. TACT downsamples the data sequences without affecting the shape of phase waveforms. The computation time for DTW can be reduced by  $4\times$  if we downsample each phase sequence by  $2\times$ . In our experiment, as the sampling rate of a reader can be up to  $80\text{Hz}$  (i.e., 80 readings of phase values per second), we can downsample to  $40\text{Hz}$  or a lower reading rate in practice. Note that a person usually performs an activity at a speed of less than  $3\text{m/s}$ . The corresponding frequency is  $18.5\text{Hz}$  ( $3\text{m/s} \times 2/\lambda$ ). If we sample the phase data at a sampling rate larger than  $37\text{Hz}$ , the phase signal can be recovered without any loss according to the Nyquist Theorem.

Fig.8(a) shows the warping results of the two phase waveforms of walking activity. We observe that for better alignment, one data point in a phase waveform is matched against one or multiple points in the other phase waveform. The corresponding peaks and valleys of the two waveforms are properly aligned. DTW finds the optimal alignment (depicted as a green line in Fig.8(b)) by exploring the cost matrix whose overall distance is minimized.

**3.2.2 Activity Moving Speed.** Human activities influence the phase of wireless signal. A faster body movement results in more rapid fluctuation of phase waveform. Therefore, the speed of an activity can be inferred from the frequency of phase waveform. Based on our proposed model in Section 2.2, the phase waveform of the combined signal have the same period as the dynamic signal waveform when an object moves continuously in a relatively static environment. Note that commodity RFID readers only report the phase values of the combined signal, while the dynamic signal reflected from human body cannot be directly measured. We observe that the two signals share the same period, which indicates that they have the same frequency. Thus, we can directly measure the frequency of combined phase waveform to infer the moving speeds of activities.

Simply using FFT to extract frequency of the phase waveform loses the temporal information and cannot reveal the instantaneous speed of the activity. TACT exploits Short-Time Fourier Transform (STFT) to calculate the moving speed. STFT divides the signal into multiple short segments with the same length and performs FFT on each segment such that the temporal information is preserved and the instantaneous frequency of the signal can be extracted. When performing STFT, different types of window functions result in various levels of frequency distortion suppression caused by spectral leakage. We choose Hamming window to effectively reduce the spectral leakage, which is proven to be more suitable for spectral leakage suppression in a signal with multiple frequency components [8].

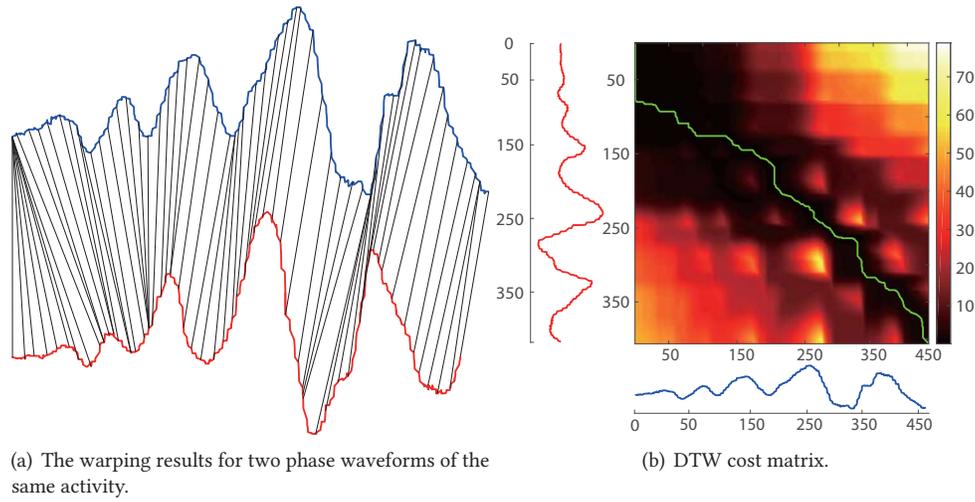


Fig. 8. Dynamic Time Warping for phase waveform comparison.

In our experiment, we use the Impinj Speedway R420 RFID reader. The commodity reader does not support a fixed read rate irrespective of the number of tags in the environment. We use the AutoSet Dense mode of the reader since in this mode the reader can adjust automatically and dynamically according to the environment and measured throughput. The read rate in our paper is calculated by averaging the number of readings for a certain time period. In our experiment, the read rate turns out to be approximately 80 readings per second. Different window sizes introduce different frequency and temporal resolution. We set the window size to 128, which provide a frequency resolution of 0.625Hz. This corresponds to an activity speed of 0.1m/s with the wavelength of 0.325m, which is sufficient for catching human activities such as walking and falling down. Furthermore, we apply an overlapping ratio of 75% to reuse parts of adjacent past data to improve temporal resolution by 4 $\times$  when performing FFT on each segment.

Note that although these parameters are empirically set, all parameters are pertinent to physical meanings. The parameters are set to satisfy the physical properties of human activities such that the performance of TACT can be maximized. More importantly, users or developers can manually tune or design parameter optimization algorithms to automatically adapt the parameter settings at different environments.

During a short time interval between two consecutive phase readings, the propagation length changes approximately 2 times faster than the body movement [26, 30]. For instance, when a user moves from  $A$  to  $B$  in Fig.4(a),  $AB \approx (d_2 - d_1)/2$ . Thus, we estimate the instantaneous activity moving speed as  $v = f \times \lambda/2$ , where  $f$  is the instantaneous frequency of the phase waveform and the wavelength  $\lambda = 0.325mm$ . In Fig.9, the spectrogram illustrates the frequencies of the phase waveform. In Fig.9(a), we can clearly observe that the high-frequency band  $f$  (e.g., red grids) is intensively distributed around 4 ~ 5Hz throughout the walking period, which indicates a speed of 0.6 ~ 0.75m/sec. While in the falling activity the frequency band suddenly increases from 5Hz to 8Hz at around 0.3 – 0.7s, as shown in Fig.9(b), which reveals a sudden acceleration from 0.8 – 1.3m/sec.

**3.2.3 Activity Moving Distance.** We extract moving distance from the signal phase values of combined signal. Note that the phase of combined signal is dominated by the phase change of dynamic component, as illustrated

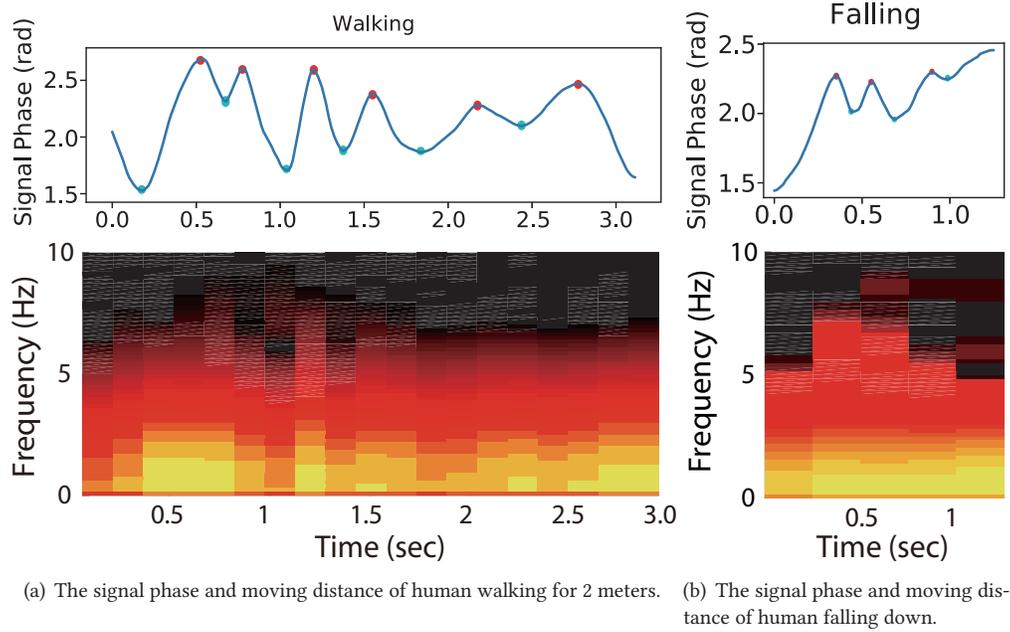


Fig. 9. Waveforms and spectrograms for walking and falling down.

in Fig.4(b). The phase range of the dynamic component is  $[0, 2\pi]$ , while the phase range of combined signal is  $[\theta, \Delta\theta + \theta]$ . Thus, the propagation distance change  $D$  can be expressed as follow:

$$D \approx \sum_{i=1}^{N-1} \frac{\lambda}{2\Delta\theta} \times (|\theta_{i+1} - \theta_i|) \quad (3)$$

where  $N$  is the total number of sampling points,  $\lambda$  is the wavelength,  $\theta_{i+1}$  and  $\theta_i$  are two successive sampling phase values of the combined signal, which can be directly measured.

In our analysis, we assumed that signal strength of dynamic component remains constant. In practice, however, the strength may change over time. As a result, the amplitude of the phase waveform may change, as shown in Fig.9. To measure  $\Delta\theta$  in this case, we find all the local maximum and minimum values illustrated with red dots and green dots in Fig.9, respectively. Then we dynamically track and update  $\Delta\theta$  as the difference between the peak and the valley in each segment when applying Eq.(3).

### 3.3 Classification

In model training stage, we combine the 3 coarse-grained features (e.g., speed, distance and time duration), the fine-grained feature of phase waveforms and three other commonly used features that can reflect the data central tendency in statistic including mean, median and standard deviation to train classifiers. In classification stage, we extract the corresponding features from the signal phase readings of unknown activities, and use the trained classifiers to recognize unknown activities. We implement and evaluate multiple classifiers (e.g., Neural Networks, Decision Tree, Support Vector Machine, Random Forest, Naive Bayes and Quadratic Discriminant Analysis).

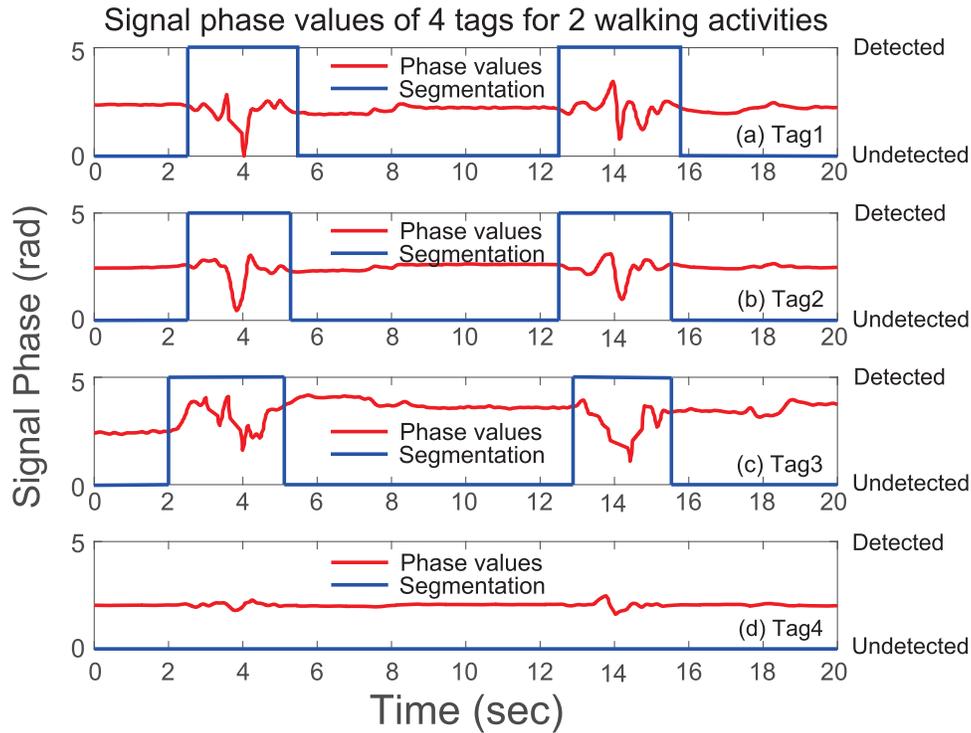


Fig. 10. The simultaneous phase readings of 4 tags when performing 2 walking activities.

In our work, all the classifiers are trained with the same 14-D feature vector labelled with ground truth activities. Among the total 14 dimensions, there are 8 fine-grained features of DTW distances, which measure the DTW distances between the phase waveform of an unknown activity and 8 template phase waveforms of 8 different activities. Along with the 8 fine-grained features of DTW distances, the 14-dimension vector also contains 3 coarse-grained features (i.e., speed, distance and duration) and 3 statistic measurements (mean, median and standard deviation). The output of classifiers for each input data is a 8-D vector which measures the probabilities of 8 possible activities. If only one tag is used in the activity recognition, the highest probability indicates the activity. If multiple tags are used, we combine the results of the multiple tags to enhance the robustness.

### 3.4 Leveraging Multiple Tags

Objects in the environment sometimes cause occasional blockage of the backscatter path, which induces low reading rate of a tag. Besides, the communication range of RFID systems is limited to approximate 10m, and the reading rate of the tag decreases as the communication distance increases. Intuitively, deploying multiple tags in the environment can alleviate blockage of the tags and can increase the coverage area. More importantly, the marginal cost of deploying extra tags is very low compared with other radio-based approaches (e.g., WiFi, Doppler radars). As such, we leverage multiple tags deployed in the environment to enhance our system.

Fig.10 shows the signal phase values of 4 tags in the environment to simultaneously detect the same walking activities for 20s. During this period, a volunteer walks for around 3s, stands for a while, and walks for another 3s, in the area covered by the 4 tags. The tags are placed in the environment with different distances to the reader.

We observe that the phase waveforms of all 4 tags fluctuate when the volunteer walks while remain stable when the volunteer stands still. However, the phase waveform of Tag4 experiences less fluctuation than the other three tags (Tag1, Tag2, and Tag3). Our segmentation algorithm may not be able to detect the activity if only Tag4 is deployed. By deploying multiple tags, the phase waveforms of the other 3 tags can be detected and segmented (depicted as the blue line in Fig.10) for further processing.

TACT merges the results of  $n$  tags at the classification step. Specifically, for data sequence of each tag  $j$  which is successfully segmented by our segmentation algorithm, TACT calculates the probability  $p_j(Y = c_k|X)$  of each activity that the inputted phase sequence  $X$  belongs to using our trained classifier. Then, the recognition result is determined by the activity  $c_k$  with the maximal average probability:

$$\arg \max_{c_k} \sum_{j=1}^n p_j(Y = c_k|X)/n \quad (4)$$

where  $c_k \in \{activity_1, activity_2, \dots, activity_k\}$  and  $n$  is the number of the segmented tags. Note that segmented phase waveforms can be processed in parallel to reduce the response time.

## 4 EVALUATION

### 4.1 Experiment Settings

**Hardware configuration:** Impinj Speedway R420 RFID reader has been widely used in RFID community since it provides accurate phase measurement results and can achieve very high localization and tracking performance. We notice that although various commodity readers (e.g., other models from Impinj, Alien RFID reader) also provide accurate phase values, in practice some commodity readers output incorrect phase readings. Therefore, we implement our system using the Impinj Speedway R420 RFID reader and commodity passive tags (Fig.11(a)) without any hardware or firmware modification. The antenna of the reader can transmit and receive the RF signal simultaneously and the communication range is up to 10m. The reading performance however decreases as the communication range increases. The channel frequency hops amongst 10 channels ranging from 920.25 – 924.75MHz. In data preprocessing, the phase readings collected over different frequency channels are mapped to a fixed channel (i.e., 920.25MHz with corresponding wavelength  $\lambda = 32.5cm$ ). The commodity reader sends the received phase readings via Ethernet cable to a Thinkpad laptop with Intel Core I7-6600 CPU and 12GB RAM. The data processing algorithms are implemented in Python and Matlab. We use Scikit-learn [21] to classify different activities. Scikit-learn is a machine learning library for the Python programming language, which includes various classification, regression and clustering algorithms. Most mainstream classifiers are packaged in Scikit-learn library for easy use. In the evaluation, we use the library to build different types of classifiers and conduct performance evaluation. We mainly adopt the default parameter settings in Scikit-learn library.

**Participant statistics:** We recruit 12 volunteers (5 females) to stand between the antenna and the tag array one by one and perform 8 different types of activities in 3 different meeting rooms covered by the reader at different times across 3 months. These volunteers are between 24 and 58 years old ( $\mu = 32.8$ ), weighed between 45 and 80 kg ( $\mu = 66.7$ ), and are between 155 and 182 cm tall ( $\mu = 171$ ). During the experiments, the volunteers wear their daily attire with different fabric materials and perform the activities at their comfortable speeds. The participants perform only one activity at a time. We do not aim to recognize multiple activities simultaneously performed (e.g., standing up while raising one hand, raising one hand while dropping other other hand).

**Experiment environment:** The size of the rooms are  $2m \times 3m$  for roomA,  $6m \times 8m$  for roomB and  $8m \times 10m$  for roomC, respectively as illustrated in Fig.11(b)-(d). The blue stars and the red dots show the placement of reader's antenna and tag array, respectively. There are multiple chairs, desks, computers and projectors placed in the rooms. Due to the daily usage, the arrangement of furniture in each room may change during the experiment

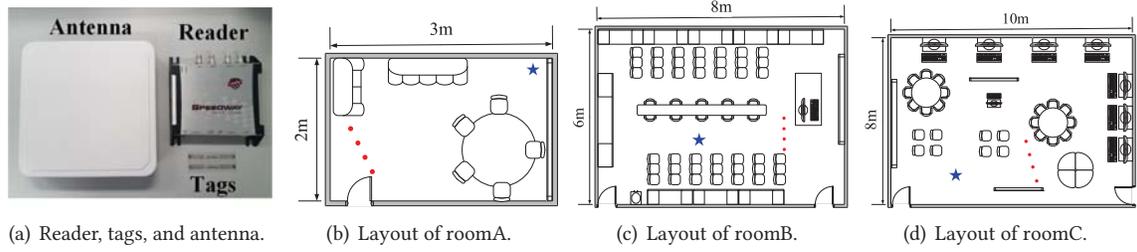


Fig. 11. Commodity Impinj RFID system used in our TACT and three different environmental layouts .

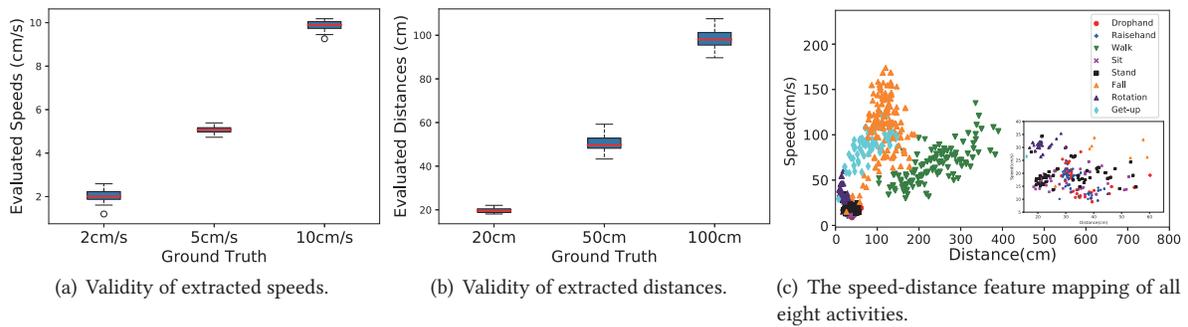


Fig. 12. Evaluation of extracted features.

period. The volunteers perform different activities multiple times nearby the RFID system during each experiment period. In most experiment settings, there are LOS paths between tags and readers. In practice, the LOS paths can be easily created by deploying more tags in the environment. As such, even if some LOS paths between reader and tags are blocked, other LOS paths can be used for activity recognition.

**Data collection and analysis:** The segmentation algorithm automatically crops candidate phase segments. If the segmentation algorithm cannot detect the variances of phase and miss the activities, we consider those cases as detection errors. All the phase segments are manually labeled with ground truth activities for training and testing. We collect 1920 signal phase sequences of eight activities. Each activity is repeated 240 times by different volunteers in 3 different rooms at different times. These activities are in random directions within the monitoring area at different speeds and distances. In total, we use 640 traces for training and 1280 traces for evaluation. In our evaluation, we use 10-fold cross-validation.

#### 4.2 Evaluation of Different Features

During the experiment, the ground truth of speeds and distances for different activities are generally hard to measure and record in practice. This is because when a user performs different activities, multiple body parts will simultaneously reflect RFID signals. Even if we regard a user as a single reflection point, it is still challenging to measure the ground truth of moving speeds or distances, since a user’s movement speeds may vary during every measurement interval.

Instead, we conduct several experiments to evaluate the accuracy of the extracted features. Note that the coarse-grained features are extracted from the proposed signal reflection model. Therefore, to evaluate the accuracy of the extracted coarse-grained features, we start to test the validity of our signal reflection model. We

move a metal reflector along the perpendicular bisector between the reader and tag, in which the ground truth of the speeds and distances can be controlled and accurately measured. To test the validity of moving speeds, we fix the moving distance to 100cm and move the reflector at constant speeds of 2cm/s, 5cm/s and 10cm/s, respectively. On the other hand, we fix the moving speeds at 5cm/s and move the reflector at distances of 20cm, 50cm and 100cm, respectively. For each speed and distance, we repeat the experiment 10 times. In Fig.12(a) and Fig.12(b), we observe that the extracted speeds and distances almost match the ground truth (i.e., read lines) with small biases.

In the following, we evaluate the speeds and distances of human activities. We combine the quantified speed and distance together from all our experiments conducted in different rooms at different times. Fig.12(c) shows the estimated speeds and distances for different activities. Each point in Fig.12(c) represents a measured *distance* and *speed* of an activity. The speeds of falling and walking significantly differ with other activities, which are around  $0.8 \sim 1.5m/s$  for falling and  $0.4 \sim 1.0m/s$  for walking, respectively. Most of the speeds of walking are slower than falling activities. In addition, from the measured distances of different activities, we observe that walking and falling down have longer moving distances, which are around  $1 \sim 2m$  for falling and  $1.5 \sim 3m$  for walking, while most of walking distances are longer than falling. Moreover, the speeds of getting-up are relatively slower than the falling down, while the distance of these two activities has part of overlap. The distances of getting-up are shorter than walking as well.

Other activities are difficult to distinguish since they share similar speeds and distances. For example, dropping a hand and sitting down usually have similar moving speeds and most sitting down and standing up activities exhibit similar moving distances. We observe an overlap of features for those activities (depicted at the left bottom in Fig.12(c)). When we zoom in, we find that distances of drophand, raisehand, sit and stand are around  $20 \sim 60cm$  with the speeds ranging from  $10 \sim 35cm/s$ . The body rotation activity exhibits shortest distances while the speeds are relatively faster than other four activities.

We note that the measured speed and distance may be biased from the true speed and distance of a user. However, as illustrated in Fig.12(c), the measured speed and distance are still effective to differentiate some activities (e.g., 'Fall', 'Walk'). The correlated speeds and distances extracted from RF signal phase values reveal the validity of our proposed reflection model.

### 4.3 Improvement of Performance

We compare our work to the RF-Care system [41], which uses RSSI to recognize human activities. In RF-Care system, the user performs activities in front of an RFID tag array deployed as a square. In our experiment, we use collected RSSIs and implement RF-Care algorithms. All parameters in RF-Care are configured and optimized according to [41] to achieve its best performance in our experiments. In TACT system, we apply all the techniques in this paper. In this experiment, we fixed the number of tags to 4 and the communication distance to 3m. The experiment is conducted in room B. The comparison result is shown in Fig.13.

In Fig.13, we observe that TACT outperforms RF-Care, especially in identifying Raisehand, Drophand and Rotation activities. This is because the body movement of these activities are relatively small, which may not be captured with RSSI. In comparison, with our proposed model and phase measurements, the performance of each activity exceeds 89% due to reliable features and more sensitive phase values.

### 4.4 Evaluation of Different Classifiers

We evaluate the classification performance of 6 mainstream classifiers by training the same features extracted using our reflection model, including Random Forest (RF), Multi-layer Perceptron based Neural Network (NN), Decision Tree (DT), Support Vector Machine (SVM), Naive Bayes (NB) and Quadratic Discriminant Analysis (QDA). All the classifiers are implemented using Scikit-learn with their default parameter settings. For example, in SVM, the default C value is set to 1.0. In RF and DT, the minimum number of samples required to split an

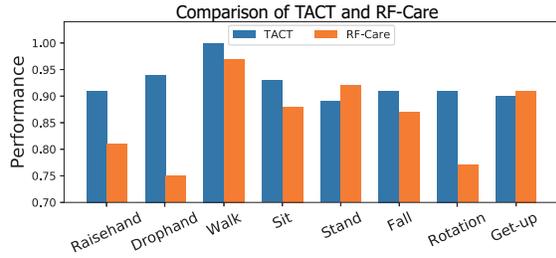


Fig. 13. The comparison of performance with existing work

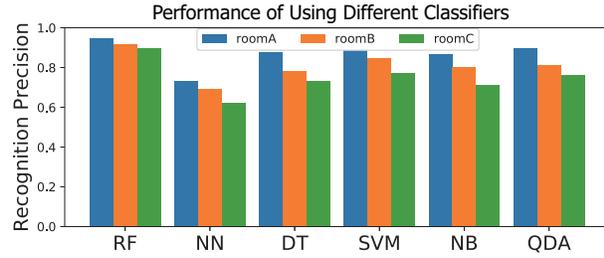


Fig. 14. The comparison of performance with different classifiers.

Table 1. Performance of TACT

	Raisehand	Drophanhand	Walk	Sit	Stand	Fall	Rotation	Get-up	Non-activity
precision	0.93	0.90	0.98	0.94	0.93	0.88	0.97	0.93	0.95
recall	0.91	0.94	1.00	0.93	0.89	0.91	0.91	0.90	0.92
f1_score	0.94	0.92	0.99	0.93	0.91	0.89	0.94	0.91	0.93

internal node and to be at a leaf node is set to 2 and 1, respectively. The minimal number of trees in RF is 2. With the same training data set, we vary the number of layers ranging from 2 ~ 6 to optimize the performance of NN. We find that 3 layers with 100, 50, and 20 neurons in each layer achieve the best performance. The ReLU function is used as activation function and the learning rate is set to 0.001. We evaluate the performance of each classifier in different rooms, as shown in Fig.14. From Fig.14, we observe that Random Forest achieves the highest precision of average precision of 93.5% in all three rooms, which is 10% higher than QDA (the second best method). The Random Forest method can deal with anti-overfitting problem when the data size is relatively small. Limited by the training data set, NN has the worst performance. Other classifiers achieve an average precision around 80%.

We observe that the performances decrease when the size of room increases. This is because a larger room with more complex layout contains more signal reflections in the environment. The performance of the Random Forest method decreases relatively slower than other methods since Random Forest method is less susceptible to the noise in the data. Thus, our work adopts the Random Forest method as the classifier in the activity recognition. In the following, we report the performance of TACT with the Random Forest method. We leave the fusion of multiple classifiers as future work.

#### 4.5 Evaluation of System Performance

We evaluate the activity recognition performance of TACT and show the results using confusion matrix in Fig.15. We calculate the recall (i.e.,  $\frac{TP}{TP+FN}$ ), precision (i.e.,  $\frac{TP}{TP+FP}$ ), and f1\_score ( $\frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$ ) of each activity based on the data provided in the confusion matrix, where  $TP$  is the true positives,  $FN$  is the false negatives, and  $FP$  is the false positives, respectively.

We evaluate TACT by deploying 4 tags in 3 different rooms with communication distances varying from 2 ~ 4m. The 4 tags are placed along a line and each tag is separated by approximately 0.5m. We make sure that the 4 tags can be read with the reader after deployment. In this evaluation, we choose Random Forest classifier since its performance exceeds other classifiers. Fig.15 depicts the confusion matrix of average recognition results of TACT, where each row represents the actual activity while each column shows the activities predicted by TACT. We compare the recognition precision of TACT to the precision when only using coarse-grained features Fig.15(a) and only using fine-grained phase waveforms Fig.15(b). In the first scenario, we choose 3 coarse-grained

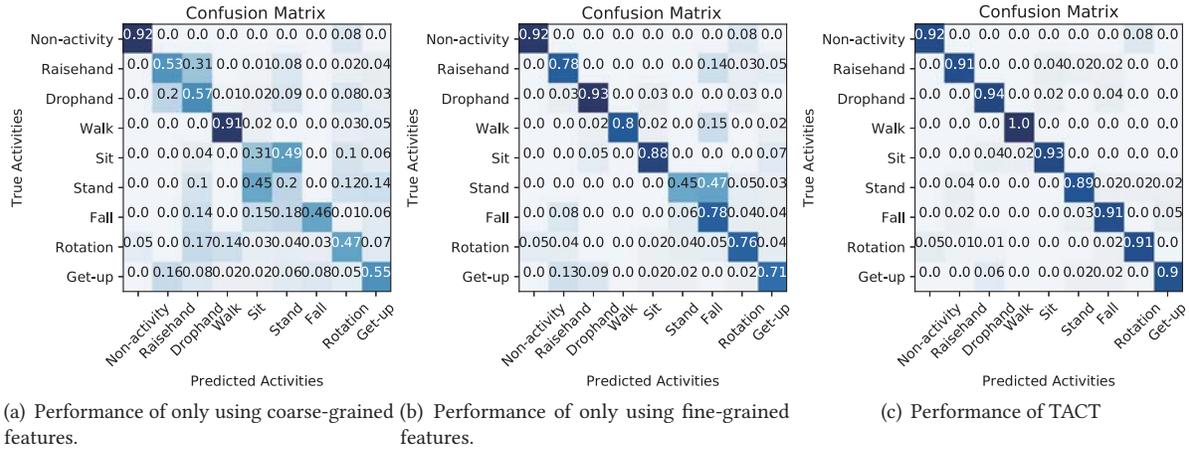


Fig. 15. Contribution of coarse-grained features and fine-grained features.

features (e.g., speed, distance and time duration of activity) plus central tendency features. In the second scenario, we only examine the phase waveforms with DTW plus central tendency features. In Fig. 15(a), we notice that when only coarse-grained features are applied, the recognition precision of raise-hand, drop-hand, sit, and stand is very low because these four activities usually have similar duration, speed, and distance, while the precision of walk activities are over 80% due to their different speed, distance and time duration compared to other activities. In Fig. 15(b), when we only examine the fine-grained phase waveforms, the precision of drop-hand and sit activities achieve over 88% due to their distinct shapes of phase waveform. However, 47% of stand activities are wrongly classified as fall since these two activities share the similar phase waveforms. In order to achieve the best performance, we combine the coarse-grained and the fine-grained features. Fig. 15(c) and Table 1 show the performance of TACT. In Fig. 15(c), we find that the performance of each activity exceeds 89%. Our data segmentation method correctly captures most of activities. Within all activities, 5% of rotation activities are not captured since the rotation activity has relatively small body movement and may not cause obvious phase change. Other activities are successfully detected due to their large body movement. In Table 1, the precision, recall, and f1\_scores are approximately 0.9 for each activity.

We summarize the results according to different activities and calculate the average precision to quantify the overall performance as well. By combing all features, TACT achieves the average precision of 93.5%. The improvement of TACT is mainly because the coarse-grained features and the fine-grained features complement in differentiating different activities.

#### 4.6 Impact of Communication Distances and Number of Tags

We evaluate the performance of TACT at different distances between reader and tag with different number of tags deployed in the same environment in Fig. 16. The number of tags varies from 1 ~ 4 and the distance varies from 2m to 4m. We place the tags along a straight line, which are separated by 0.5m. We observe that the performance of TACT increases with the number of tags increases for all scenarios with different communication distances. This is because phase readings from multiple tags can provide more reliable information than a single tag does, especially when under some harsh environments. In addition, our segmentation algorithm neglects invalid phase readings and our merging algorithm for multiple tags can effectively combine the recognition results of multiple tags to achieve higher precision.

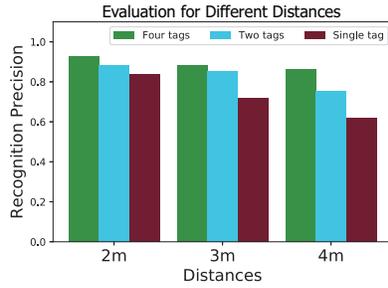


Fig. 16. Evaluation for different communication distances using different number of tags.

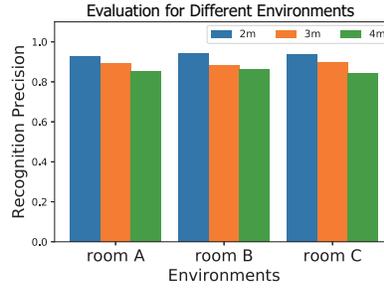


Fig. 17. Evaluation for different environments using different communication ranges.

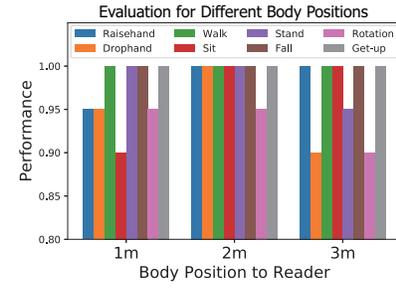


Fig. 18. Evaluation for different body positions.

Note that when the distance increases, the reader receives weaker backscatter signals and affects recognition precision. Thus, the precision of TACT decreases as the communication range increases when using different number of tags. We notice that the average precision only decreases 6% from 92% ~ 86% when using 4 tags at different communication ranges, while decreases 21% from 83% ~ 62% for only using one tag. The experiment results indicate that we can deploy more tags to improve the activity recognition performance.

#### 4.7 Impact of Different Environments

We test our TACT using the data all collected in the same room. For each room, we use the phase sequences from one experiment for training and the rest two experiment for testing. We test TACT in 3 different rooms with different sizes and environments. We collected 640 activity corresponded phase sequences in each room from different volunteers at 3 different times. The layout of each room did not keep the same for each experiment. In this evaluation, we fix the number of tags to 4, which are deployed along a straight line as well with the same distance of 0.5m between every two adjacent tags. In each room, we evaluate TACT with the communication distance varying from 2m to 4m. TACT is trained and evaluated using the features extracted in the same room.

Fig. 17 shows the performance of TACT in different environments under different communication ranges. We observe that the average precision achieves 93.5% when the communication range is 2m. Although the average precision slightly decreases as the communication distance increases, it still remains 85.5% as the communication range reaches 4m. The experiment results reveal that TACT is able to work in different environments with consistent recognition precision.

#### 4.8 Impact of Different Body Positions

In training stage of our experiment, the training data contains activities performed and captured at different positions within the communication range. The collected training data contains signal phase values introduced by the activities at different positions. As a result, our trained model successfully classifies the activities even performed at different positions. In evaluation stage, we also ask the volunteers to perform activities within the communication range at random places. The average performance of TACT can be up to 93.5%.

To further evaluate the trained model when people are performing activities at different positions, we ask two volunteers to perform activities at the distance of 1m, 2m, and 3m to the reader along with the LOS path, respectively. The distance between the reader and tag is set to 4m, and the number of tags is fixed to 4. Both the reader and tags are mounted on tripods at 1m height. The experiment is conducted in room B. In this case, we collect 480 traces for testing. The experiment result is shown in Fig. 18.

In Fig.18, we observe that almost all the activities are correctly classified by our trained model even if the activities are performed at different positions. We find that, the system achieves higher performance when users perform activities near the middle area between reader and tag. The evaluation results reveal that TACT can deal with the activities performed at different positions within the communication range.

#### 4.9 Limitations and Future Work

**Model Generalization to different users:** In our study, we apply 10-fold cross-validation to evaluate TACT. Although the differences in phase waveforms for the same activities performed by different people can be mitigated by applying DTW, capturing sufficient variations of phase waveforms generally needs a large number of volunteers. Collecting such a number of training data is inconvenient and sometimes impractical. One possible solution is to use data augmentation techniques [42] to imitate the variations of phase waveforms based on the real collected data. For example, we may perform horizontal shrinkage and expansion on phase waveforms to imitate the activities performed at different speed. The augmented data may contain sufficient variations of activities and provide more reliable information for training the model.

**Model Generalization to different environments:** Although the overall performance of TACT achieves 93.5% in three different rooms, TACT is trained and evaluated using the features extracted in the same room. Currently, our trained model in one room cannot be directly used in different rooms due to different environments. As a result, one needs to re-train the model to learn the background information of a new environment. We plan to apply transfer learning techniques such that the knowledge of model trained in one room can be transferred to other rooms without re-training a new model from scratch.

**Detection of Multiple Persons:** Our current design of TACT cannot identify activities simultaneously performed by multiple persons. In future work, we plan to deploy more readers with multiple antennas and more tags in the environment and apply multiple source separation techniques to select the most relevant tags that are only affected by one person.

## 5 RELATED WORK

RF-based detection systems have attracted plenty of interests in recent years. RFID readers can be mounted on human hands to capture object use by detecting the presence of a tagged object [7]. In [29], the authors use passive tags to extract both spatial and temporal features to characterize various activities. RF-IDraw [28] exploits antenna array and beam steering technique to track the trajectory of a labeled object. PinIt [27] measures the angle of arrival of reflected signals collected from multipath profiles to locate the tags. Tagoram [39] locates objects by tracking RFID tags by analyzing phase values collected using commodity RFID systems. By tracking the movement of RFID tags attached to the dumbbells, FEMO [5] provides an integrated free-weight exercise monitoring system. TagBreathe [9] monitors human breaths for multiple persons. These works need the users to carry devices such as sensor motes and RFID reader motes. Our work differs from these works in that it does not require users to carry any devices. IDSense [13] aims to build a human object interaction detection system by measuring the changes in the physical layer of the communication channel between the RFID tag and reader. This work uses RFID tags to label daily items (e.g., cloth, glasses, books), and detection human object interaction by measuring the changes in RSSI and RF phase when users directly interact with those RFID-labelled items. Compared with this work, our work aims to leverage all deployed RFID tags in the environments.

Contact-free systems have also been investigated in recent years. Tadar [40] is able to track moving objects behind a wall. RFIPad [4] enables in-air handwriting without requiring to carry any device. These works restrict users to a small area to perform gestures (e.g., within several decimeters behind tag arrays) or need well-designed signal amplifiers to penetrate a concrete wall. In contrast, our TACT system does not need extra devices and can monitor human activities within multiple meters sensing range. RF-Care [41] is a pioneering work which aims to

build a device-free posture recognition for elderly people using passive RFID arrays. This work detects human activities by analyzing the RSSI patterns from an RFID tag array. Our work differs from RF-Care in that it builds a reflection model and uses phase measurement of the reflection signal, which provides better resolution. Some deep learning based methods are proposed to automatically identify human activities [6, 14]. However, such methods need a large number of labeled data to train neural network models [14]. Moreover, such methods are susceptible to the environmental change such as arrangement of furniture in the environments [6]. In contrast, our TACT system is more resilient to the environmental change since TACT is able to extract the dynamic components due to human activities and the static components reflected from background can be eliminated.

WiFi-based tracking and activity recognition systems have been proposed in previous works [34, 36, 37]. E-eyes [32] uses fine-grained CSI to identify human activity patterns. WiSee [22] and Allsee [12] can recognize multiple home gestures. CARM [30] is capable of detecting human activities by quantifying the speed of motion and CSI power. WiHear [25] is capable of tracking mouth motion when people are talking. WiSleep [15] uses commodity WiFi devices to detect human respiration when people are sleeping. Phasebeat [31] is able to monitor human breathing and heartbeat with commodity WiFi devices by analyzing CSI.

WiTrack [1] is a contact-free RF-based motion tracking system, which measures distance to human body using FMCW technique. Vital Radio [2] uses FMCW system to track human breathing and heart rate without body contact. WiSprio [19] uses Doppler Radar to track human breaths during sleeping. Ultrasonic sensor [18, 43], laser sensor [11], and wearable sensor [35] are also used to detect human activities. However, these techniques need either direct contact to human body [35] or specialized hardware [11, 18], which are not readily available on the market and incur high cost for civilian use and large-scale deployment. In contrast, we use commodity systems to build a contact-free activity recognition system.

## 6 CONCLUSION

In this paper, we propose a contact-free activity recognition system using commodity RFID systems. The major contribution of this paper is to build a contact-free model which can be used to quantify the correlation between signal phase values and key features of human activities. We carefully select the features by combining coarse-grained features extracted from the proposed model with fine-grained phase waveform based features to differentiate different human activities. We have implemented the contact-free activity recognition system with a COTS RFID system and evaluated its performance at different communication distances, with different number of tags and in different environments. We invited 12 participants to test and evaluate our system in various scenarios. We carefully evaluate error distributions of each activity and present comprehensive evaluation results. The experiment results show that TACT is robust under different experimental settings and can achieve the average recognition precision of up to 93.5%.

## ACKNOWLEDGMENTS

The work is supported by the National Nature Science Foundation of China (No. 61702437) and Hong Kong ECS under Grant PolyU 252053/15E. We are grateful to reviewers for their insightful comments and help.

## REFERENCES

- [1] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D Tracking via Body Radio Reflections. In *USENIX NSDI*.
- [2] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *ACM CHI*.
- [3] Centers for Disease Control and Prevention. [n. d.]. Important Facts about Falls. <https://www.cdc.gov>.
- [4] Han Ding, CHEN Qian, Jinsong Han, Ge Wang, Wei Xi, Kun Zhao, and Jizhong Zhao. 2017. RFIPad: Enabling Cost-efficient and Device-free In-air Handwriting using Passive Tags. In *IEEE ICDCS*.
- [5] Han Ding, Longfei Shangguan, Zheng Yang, Jinsong Han, Zimu Zhou, Panlong Yang, Wei Xi, and Jizhong Zhao. 2015. Femo: A platform for free-weight exercise monitoring with rfids. In *ACM SenSys*.

- [6] Xiaoyi Fan, Wei Gong, and Jiangchuan Liu. 2018. TagFree Activity Identification with RFIDs. In *ACM UbiComp*.
- [7] Tao Gu, Liang Wang, Hanhua Chen, Xianping Tao, and Jian Lu. 2011. Recognizing multiuser activities using wireless body sensor networks. *IEEE transactions on mobile computing* 10, 11 (2011), 1618–1631.
- [8] Fredric J Harris. 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* 66, 1 (1978), 51–83.
- [9] Yuxiao Hou, Yanwen Wang, and Yuanqing Zheng. 2017. TagBreathe: Monitor Breathing with Commodity RFID Systems. In *IEEE ICDCS*.
- [10] Meng Jin, Yuan He, Xin Meng, Yilun Zheng, Dingyi Fang, and Xiaojiang Chen. 2017. FlipTracer: Practical Parallel Decoding for Backscatter Communication. In *ACM MobiCom*.
- [11] Ossi Kaltiokallio, Hüseyin Yiğitler, Riku Jäntti, and Neal Patwari. 2014. Non-invasive respiration rate monitoring using a single COTS TX-RX pair. In *IEEE IPSN*.
- [12] Bryce Kellogg, Vamsi Talla, and Shyamnath Gollakota. 2014. Bringing Gesture Recognition to All Devices. In *USENIX NSDI*.
- [13] Hanchuan Li, Can Ye, and Alanson P Sample. 2015. IDSense: A human object interaction detection system based on passive UHF RFID. In *ACM CHI*.
- [14] Xinyu Li, Yanyi Zhang, Ivan Marsic, Aleksandra Sarcevic, and Randall S Burd. 2016. Deep learning for rfid-based activity recognition. In *ACM SenSys*.
- [15] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-sleep: Contactless sleep monitoring via wifi signals. In *IEEE RTSS*.
- [16] Xiulong Liu, Keqiu Li, Geyong Min, Yanming Shen, Alex X Liu, and Wenyu Qu. 2015. Completely pinpointing the missing RFID tags in a time-efficient way. *IEEE Trans. Comput.* 64, 1 (2015), 87–96.
- [17] Xiulong Liu, Xin Xie, Keqiu Li, Bin Xiao, Jie Wu, Heng Qi, and Dawei Lu. 2017. Fast tracking the population of key tags in large-scale anonymous RFID systems. *IEEE/ACM Transactions on Networking (ToN)* 25, 1 (2017), 278–291.
- [18] Se Dong Min, Jin Kwon Kim, Hang Sik Shin, Yong Hyeon Yun, Chung Keun Lee, and MyoungHo Lee. 2010. Noncontact respiration rate measurement system using an ultrasonic proximity sensor. *IEEE Sensors Journal* 10, 11 (2010), 1732–1739.
- [19] Phuc Nguyen, Xinyu Zhang, Ann Halbower, and Tam Vu. 2016. Continuous and fine-grained breathing volume monitoring from afar using wireless signals. In *IEEE INFOCOM*.
- [20] Jiajue Ou, Mo Li, and Yuanqing Zheng. 2015. Come and be served: Parallel decoding for cots rfid tags. In *ACM MobiCom*.
- [21] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.
- [22] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *ACM MobiCom*.
- [23] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. 2015. WiBreathe: Estimating respiration rate using wireless signals in natural settings in the home. In *IEEE PerCom*.
- [24] Longfei Shangguan, Zheng Yang, Alex X Liu, Zimu Zhou, and Yunhao Liu. 2017. STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization. *IEEE/ACM Transactions on Networking* 25, 1 (2017), 596–609.
- [25] Guanhua Wang, Yongpan Zou, Zimu Zhou, Kaishun Wu, and Lionel M Ni. 2016. We can hear you with wi-fi! *IEEE Transactions on Mobile Computing* 15, 11 (2016), 2907–2920.
- [26] 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 *ACM UbiComp*.
- [27] Jue Wang and Dina Katabi. 2013. Dude, where’s my card?: RFID positioning that works with multipath and non-line of sight. In *ACM SIGCOMM*.
- [28] Jue Wang, Deepak Vasisht, and Dina Katabi. 2014. RF-IDraw: virtual touch screen in the air using RF signals. In *ACM SIGCOMM*.
- [29] Liang Wang, Tao Gu, Xianping Tao, and Jian Lu. 2017. Toward a wearable RFID system for real-time activity recognition using radio patterns. *IEEE Transactions on Mobile Computing* 16, 1 (2017), 228–242.
- [30] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *ACM MobiCom*.
- [31] Xuyu Wang, Yang Chao, and Shiwen Mao. 2017. PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices. In *IEEE ICDCS*.
- [32] 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 *ACM MobiCom*.
- [33] Teng Wei and Xinyu Zhang. 2016. Gyro in the air: tracking 3D orientation of batteryless internet-of-things. In *ACM MobiCom*.
- [34] Chenshu Wu, Zheng Yang, and Chaowei Xiao. 2018. Automatic radio map adaptation for indoor localization using smartphones. *IEEE Transactions on Mobile Computing* 17, 3 (2018), 517–528.
- [35] Lei Xie, Xu Dong, Wei Wang, and Dawei Huang. 2017. Meta-Activity Recognition: A Wearable Approach for Logic Cognition-based Activity Sensing. In *IEEE INFOCOM*.
- [36] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise power delay profiling with commodity wifi. In *ACM MobiCom*.
- [37] Yaxiong Xie, Yanbo Zhang, Jansen Christian Liando, and Mo Li. 2018. SWAN: Stitched Wi-Fi ANtennas. In *ACM MobiCom*.

- [38] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cekan, Yingying Chen, Marco Gruteser, and Richard P Martin. 2011. Detecting driver phone use leveraging car speakers. In *ACM MobiCom*.
- [39] Lei Yang, Yekui Chen, Xiangyang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. 2014. Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices. In *ACM MobiCom*.
- [40] Lei Yang, Qiongzhen Lin, Xiangyang Li, Tianci Liu, and Yunhao Liu. 2015. See through walls with cots rfid system!. In *ACM MobiCom*.
- [41] Lina Yao, Quan Z Sheng, Wenjie Ruan, Tao Gu, Xue Li, Nick Falkner, and Zhi Yang. 2015. Rf-care: Device-free posture recognition for elderly people using a passive rfid tag array. In *ICST MobiQuitous*.
- [42] Xiao Zeng, Kai Cao, and Mi Zhang. 2017. MobileDeepPill: A small-footprint mobile deep learning system for recognizing unconstrained pill images. In *ACM MobiSys*.
- [43] Huanle Zhang, Wan Du, Pengfei Zhou, Mo Li, and Prasant Mohapatra. 2016. DopEnc: acoustic-based encounter profiling using smartphones. In *ACM MobiCom*.

Received May 2018; revised August 2018; accepted October 2018