RF-Identity: Non-Intrusive Person Identification Based On Commodity RFID Devices

CHAO FENG, Northwest University, IoT Research Center-Northwest University, China
JIE XIONG, University of Massachusetts Amherst, USA
LIQIONG CHANG, Northwest University, International Joint Research Centre for Battery-free IoT, China
FUWEI WANG*, Northwest University, Northwest University-Jingdong Wisdom Cloud Joint Research Center for AI & IoT, China
JU WANG, IoT Research Center- Northwest University, China
DINGYI FANG, Northwest University, IoT Research Center-Northwest University, International Joint Research Centre for Battery-free IoT, China

Person identification plays a critical role in a large range of applications. Recently, RF based person identification becomes a hot research topic due to the contact-free nature of RF sensing that is particularly appealing in current COVID-19 pandemic. However, existing systems still have multiple limitations: i) heavily rely on the gait patterns of users for identification; ii) require a large amount of data to train the model and also extensive retraining for new users and iii) require a large frequency bandwidth which is not available on most commodity RF devices for static person identification. This paper proposes RF-Identity, an RFID-based identification system to address the above limitations and the contribution is threefold. First, by integrating walking pattern features with unique body shape features (e.g., height), RF-Identity achieves a high accuracy in person identification. Second, RF-Identity develops a data augmentation scheme to expand the size of the training data set, thus reducing the human effort in data collection. Third, RF-Identity utilizes the tag diversity in spatial domain to identify static users without a need of large frequency bandwidth. Extensive experiments show an identification accuracy of 94.2% and 95.9% for 50 dynamic and static users, respectively.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**.

Additional Key Words and Phrases: Person identification, RFID tag, body feature, Deep learning

ACM Reference Format:
Chao Feng, Jie Xiong, Liqiong Chang, Fuwei Wang, Ju Wang, and Dingyi Fang. 2021. RF-Identity: Non-Intrusive Person Identification Based On Commodity RFID Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 9 (March 2021), 23 pages. https://doi.org/10.1145/3448101

*This is the corresponding author

Authors' addresses: Chao Feng, Northwest University, IoT Research Center-Northwest University, China, chaofeng@stumail.nwu.edu.cn; Jie Xiong, University of Massachusetts Amherst, USA, jxiong@cs.umass.edu; Liqiong Chang, Northwest University, International Joint Research Centre for Battery-free IoT, China, clq.nwu.edu.cn; Fuwei Wang, Northwest University, Northwest University-Jingdong Wisdom Cloud Joint Research Center for AI & IoT, China, wfw.nwu.edu.cn; Ju Wang, IoT Research Center- Northwest University, China, wangju@nwu.edu.cn; Dingyi Fang, Northwest University, IoT Research Center-Northwest University, International Joint Research Centre for Battery-free IoT, China, dyf.nwu.edu.cn.

2474-9567/2021/3-ART9 $15.00
https://doi.org/10.1145/3448101

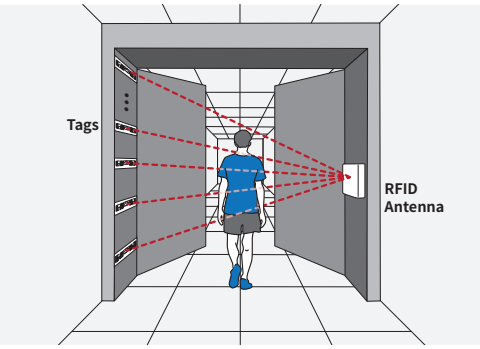


Fig. 1. The deployment of RF-Identity consists of a UHF-RFID tag array and a directional antenna.

1 INTRODUCTION

Person identification is a process of identifying an individual based on pre-recorded information. Many applications benefit from the identification of users. For example, there can be multiple occupants in a smart home environment. It is useful to know each user's identity non-intrusively to provide personalized services. For instance, when a user walks into the living room and turns on the TV, the TV can automatically switch to the user's favorite channel. The room temperature and lighting conditions can also be adjusted based on the user's preference. The system can also prevent children from operating risky home appliances such as micro-oven through user identification. Moreover, without a need for card touch, a smart door can identify and allow an authorized user to enter the room.

Traditional person identification systems use either Fingerprint [21], Voiceprint [8, 41] or camera [19, 32]. Although they have been used for decades, there are still limitations with these approaches. Camera and voice-based solutions may raise privacy concerns, while fingerprint requiring user contact is not a desirable scheme in current COVID-19 pandemic. Furthermore, these approaches still need user involvement. For example, for camera-based approaches, to make sure the camera can capture the user's face, the system usually entails the user to move multiple times physically.

Recently, RF-based person identification systems [16, 18, 52] have attracted a lot of research interests due to the appealing sensor-free and contact-free (non-intrusive) nature. Though promising, these identification systems still have multiple limitations, preventing them from being adopted in real-world applications. First, most existing systems rely only on walking patterns for user identification and assume a user keeps the same walking pattern during both training and identification phases [40, 52, 53]. This is not always true. Even the same user may exhibit large variations on his/her walking patterns due to physiological and emotional changes, resulting in identification errors. Second, existing systems require collecting a large amount of data for fine-grained person identification [11, 20, 30] and extensive retraining when new users come in, which is time-consuming and labor-intensive. Third, most existing systems require the user to move (e.g., wave the hand or walk) for identification. It is challenging for existing approaches to identify a static person. The very few approaches which are capable of identifying static users require a large frequency bandwidth [9, 13] which is not available at most commodity hardware to extract richer features.

To address the above issues, we propose RF-Identity, an RFID-based person identification system in this paper. Compared with Wi-Fi technology, Ultra-High-Frequency (UHF) RFID tags are cheap, small, and flexible for deployment. Moreover, they do not need any battery support due to the passive nature and can harvest power from the signals for backscatter transmission. The proposed system has three advantages. First, RF-Identity

combines the unique behavioral (walking pattern) features and the physiological (body shape) features of a user acquired from RFID tags, as illustrated in Fig. 1. The proposed system is therefore more robust and accurate than existing systems that rely only on the walking patterns. Second, RF-Identity can accommodate new users without retraining. Thus, it significantly reduce the amount of human effort and time cost. Third, RF-Identity can accurately identify static users without requiring a large bandwidth. To achieve these objectives, RF-Identity needs to address the following challenges.

Challenge 1: How to achieve robust person identification when a user's walking pattern varies? To tackle this challenge, we exploit multiple RFID tags to extract not just the walking pattern features but also body shape (e.g., height and width) features for identification. To obtain the body shape features, our key observation is that by carefully deploying the tags, the users with different heights and widths will cause different received signal amplitude changes. Thus, we devise a body shape extraction scheme based on the radio tomographic imaging (RTI) technique [44]. To obtain the walking pattern features, we design an attention-based module [4] combined with a residual network module to learn the walking pattern. Finally, by integrating the walking pattern and the body shape features through a hybrid deep learning framework, RF-Identity can achieve robust and accurate person identification.

Challenge 2: How to use limited amount of data to train a deep learning network for person identification and adapt it to new users without retraining? To deal with this challenge, we first design a data augmentation method to expand the size of the training data set, which dramatically reduces the data collection effort. Then, we develop a transfer learning method to adapt the trained network for new user identification. The intuition is that the neural network weights of the previously trained feature extraction module are still valid for new users. Thus, instead of retraining the whole network, we freeze the weights of the feature extraction module and only fine-tune parameters of the network's updating module (i.e., fully-connected layers) that are the key for adapting the model to new users.

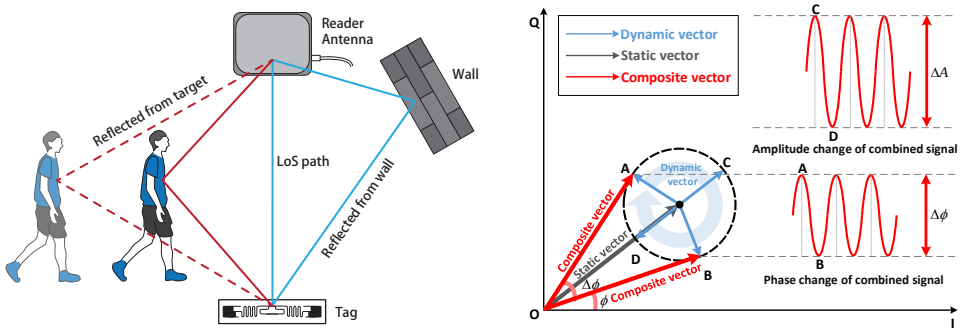
Challenge 3: How to accurately identify a static user without a large frequency bandwidth? When a user is static, we are not able to employ movement information such as walking patterns for identification. Thus, most existing systems have difficulties identifying static users [11, 16]. To address this problem, RF-Identity exploits tag diversity in spatial domain to construct phase and amplitude profiles as the static user's biometric features. Our experiment results show that different users have distinguishable phase and amplitude profiles. To achieve a high accuracy, we present a weight-based distance method to differentiate profiles among different users.

We implement our system with commodity RFID devices (12 cheap tags and one reader) and evaluate the system performance in a typical lab environment. Extensive experiments demonstrate RF-Identity can identify 50 volunteers at an average accuracy of 94.2% and 95.9% when users are walking and static, respectively.

Contribution: The main contributions of this work are as follows:

- To the best of our knowledge, this work is the first attempt to achieve person identification for both dynamic and static persons using commodity RFID devices.
- We propose a hybrid deep learning model that can efficiently utilize both walking pattern and body shape features to achieve robust and accurate dynamic person identification. We also develop a weight-based profile method to obtain rich features for static user identification.
- Extensive real-world experiments with 50 volunteers demonstrate the effectiveness and robustness of RF-Identity.

The rest of this paper is organized as follows. We first present the background knowledge in Section 2. Then Section 3 presents the system overview of RF-Identity. The details of system design are described in Section 4. The implementation and evaluation of the system are presented in Section 5 and Section 6 respectively. Section 7 discusses the limitations of RF-Identity followed by related work in Section 8. Finally, we conclude this work in Section 9.



(a) A scenario with a moving subject.

(b) The signal decomposition model.

Fig. 2. The phase change of received signal when a subject walks in the detecting area.

2 BACKGROUND

In this section, we first introduce the technical background of RFID system. Then, we conduct two preliminary studies to show the phase and amplitude variations when volunteers walk through and stand in the sensing area.

2.1 RFID Basics

An Ultra-High-Frequency (UHF) RFID system usually consists of a reader and several passive tags. The reader sends out a continuous periodic signal to activate nearby tags, and the tags change their antenna impedance to reflect signal back to the reader. The reader can then obtain the channel information including both phase and amplitude measurements [14, 42]. To understand how a user affects the received signal when a user travels through the sensing area, we consider a typical multipath indoor scenario in Fig. 2(a). We can see that the signal propagates along three paths including the line-of-sight (LoS) path, the reflection path from a wall, and the reflection path from a moving user. If we assume there are q number of reflection paths from the moving user, the received signal can be described as:

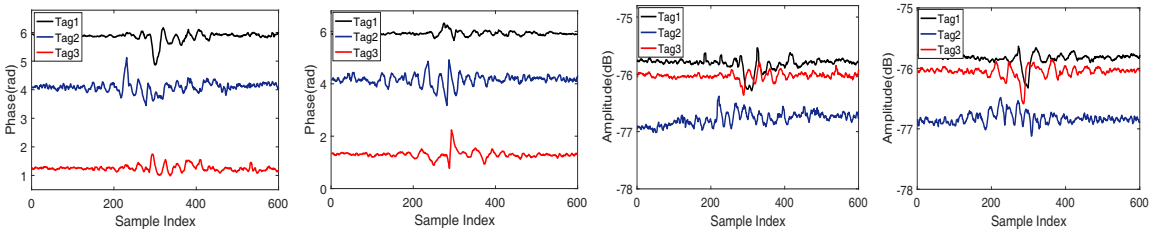
$$s(t) = \alpha_S e^{j\varphi_S} + \sum_q \alpha_q e^{j\left(\frac{2\pi}{\lambda}\int v_q(t)dt + \varphi_{dev}\right)} \tag{1}$$

where $\alpha_S e^{j\phi_S}$ is the combined complex signal containing the LoS path and static multipath, α_q is the amplitude of the $q\text{-}th$ path reflected from the human body, $v_q(t)$ is the path length change velocity corresponding to the $q\text{-}th$ path at time t, λ denotes the signal wavelength, and ϕ_{dev} is a constant phase offset induced by the tag and reader circuit.

Based on Equation (1), we find an insight that a user's presence or movement in the area of interest leads to fluctuations of the received signal. As shown in Fig. 2(b), when the dynamic component changes, the phase and amplitude of the composite component change accordingly. This implies that the phase and amplitude information of the composite signal can be used to detect a user's presence and movement. Specifically, we can exploit the phase and amplitude readings to obtain the information of the user such as the walking pattern. Furthermore, information from multiple tags can be combined to obtain richer information of the user.

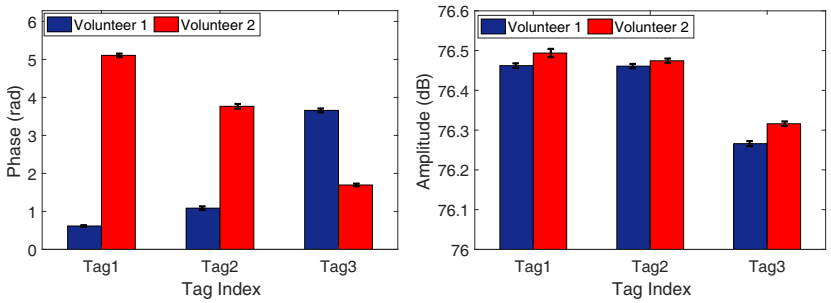
2.2 Feasibility Study and Analysis

To better understand the correlation between the signal readings and user behavioral/physiological characteristics (i.e., gait and the user body shape), we conduct two experiments at the lab entrance. Three tags and a directional



(a) Phase readings for Volunteer 1. (b) Phase readings for Volunteer 2. (c) Amplitude readings for Volunteer 1. (d) Amplitude readings for Volunteer 2.

Fig. 3. The phase and amplitude readings of Volunteer 1 and Volunteer 2 when they travel through the door attached with RFID tags.



(a) Phase. (b) Amplitude.

Fig. 4. The phase and amplitude readings vary when users stand in a fixed position facing the door attached with RFID tags.

antenna are deployed at two sides of the door. Tags are placed in a line and adjacent tags have a spacing of 0.3 m. The width of the door is 1.5 m. In the first experiment, we let two volunteers walk naturally through the sensing area one by one. Fig. 3 shows the phase and amplitude readings collected for volunteer 1 and volunteer 2. We can see that (i) different tags have different amounts of phase and amplitude changes when a human subject goes through the area. Particularly, when the subject blocks the LoS path, the variations are large; (ii) different subjects cause different amounts of changes in phase and amplitude readings. These observations imply that the phase and amplitude variations can provide us the physical and behavioral characteristics of the subject, which can be utilized for person identification.

To illustrate RF-Identity's basic idea for static person identification, in the second experiment, two volunteers are asked to stand still at a same position for a period of 10 seconds. We collect 20 measurements for each of them. Fig. 4 shows the phase and amplitude readings of each tag when different users stand at a same position. We observe that different users cause different amounts of phase and amplitude at each tag. The observation indicates that it is possible to use the phase and amplitude readings collected at multiple tags for static person identification.

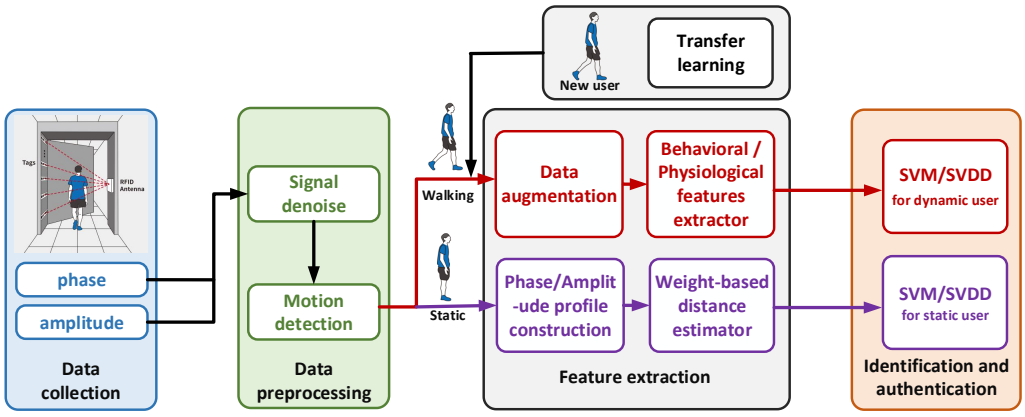


Fig. 5. System overview of RF-Identity.

3 SYSTEM OVERVIEW

RF-Identity is a low-cost person identification system built on commodity RFID devices. It involves one reader equipped with one antenna and multiple tags. The basic idea of RF-Identity is to extract unique behavioral/physiological characteristics to differentiate users when a user appears in the area of interest. The system architecture is shown in Fig. 5, which consists of the following modules:

Data Collection Module: RF-Identity collects a set of phase/amplitude readings as the baseline data when there is no user. Then, it extracts another set of readings when a user appears.

Data Pre-processing Module: Due to the environmental noise and hardware imperfection, the collected phase and amplitude readings cannot be directly employed for feature extraction. RF-Identity uses the discrete wavelet transform (DWT) algorithm to remove the environmental noise and obtain the clean phase and amplitude readings. After that, RF-Identity employs a motion segmentation scheme to segment the subject motion for feature extraction.

Feature Extraction Module: For the dynamic user (walking) case, RF-Identity first designs a data augmentation method to reduce the amount of human effort in data collection. Then a hybrid deep learning model is developed to extract and integrate the human behavioral and physiological features. Specifically, an attention-based deep learning module and a residual network are designed to obtain the human walking pattern features, and a body shape estimation module is designed to obtain the physiological features. In addition, RF-Identity introduces a transfer learning method to adapt the network for new users. In the static case, RF-Identity leverages spatially deployed tags to construct phase and amplitude profiles, and designs a weight-based method to measure the similarities of user profiles.

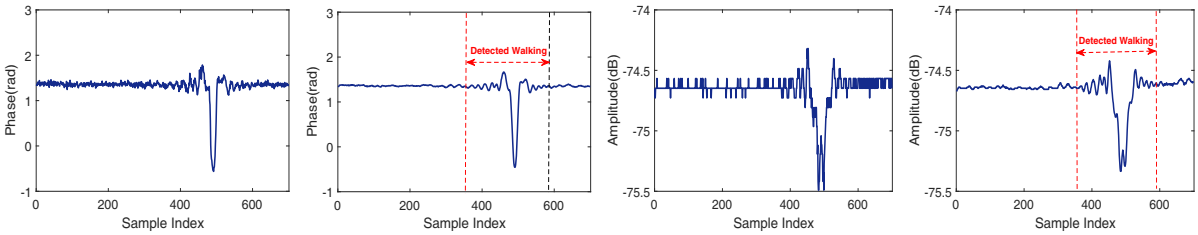
Person Identification and Authentication Module: In this module, the extracted features are fed into the SVM classifier or SVDD classifier for person identification.

4 SYSTEM DESIGN

In this section, we detail the system design of RF-Identity including data pre-processing, feature extraction, and person identification/spoof detection.

4.1 Data Pre-processing

In this sub-section, the system performs a set of pre-processing steps to obtain clean phase and amplitude readings.



(a) Raw phase readings. (b) Denoised phase readings. (c) Raw amplitude readings. (d) Denoised amplitude readings.

Fig. 6. The phase and amplitude readings.

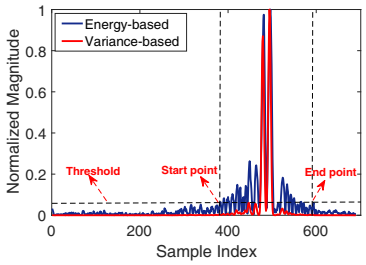


Fig. 7. The walking detection and segmentation using energy-based and variance-based sliding window.

4.1.1 Signal Denoising. Due to environmental noise and hardware imperfection, the collected phase and amplitude data are corrupted by noise (Fig. 6(a) and Fig. 6(c)). We employ the DWT algorithm [7] to help remove the noise. The key idea of the DWT filter is that the wavelet coefficient of noise is smaller than that of the signal after the raw signal (noise + signal) is decomposed by wavelet. The denoised phase and amplitude are illustrated in Fig. 6(b) and Fig. 6(d).

4.1.2 Motion Detection and Segmentation. To effectively extract the walking pattern features from the phase and amplitude measurements, RF-Identity has to perform motion detection and segmentation first. It is crucial to avoid introducing extra noise or losing useful information in the segmentation process [61] to ensure the identification accuracy in the following steps. To achieve this objective, we present an energy-based sliding window segmentation method based on the fact that the variations of phase and amplitude readings are much larger when a user enters the sensing area. We calculate the energy E as follows:

$$E_p(t) = \sum_{l=1}^{L} \phi^2(t+l) \tag{2}$$

$$E_a(t) = \sum_{l=1}^{L} A^2(t+l) \tag{3}$$

where $\phi(t)$ and $A(t)$ denote the phase reading and amplitude reading at time t, L denotes the length of the sliding window, l is the sample index. To demonstrate the effectiveness of the method, we compare it with the variance-based moving window method. Fig. 7 shows the result for the same phase stream. It is evident that the energy-based method can more accurately detect the start and end points of walking. Thus, the energy-based

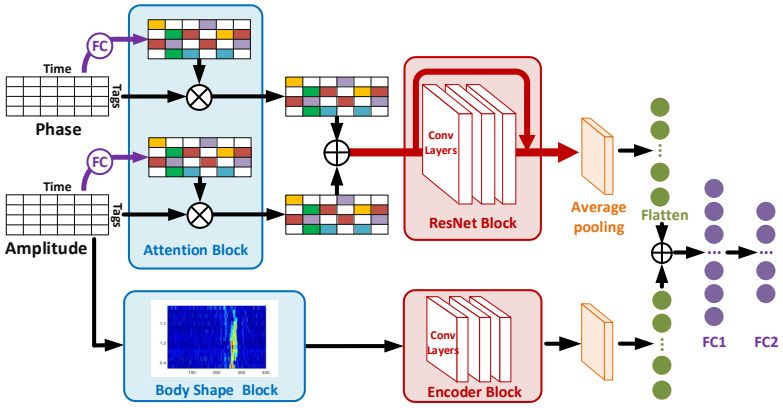


Fig. 8. The hybrid deep learning framework for dynamic user feature extraction.

method can capture the entire walking through process. In contrast, the variance-based method only detects the moment when the subject blocks the LoS path. To determine the walking interval, we incorporate both phase readings and amplitude readings to segment walking period based on Equation (2) and (3). The final start and end indexes are calculated as the average start and end index values detected from phase and amplitude streams.

4.2 Dynamic User Feature Extraction

In this subsection, our goal is to extract reliable features for dynamic person identification. Since the same subject cannot keep multiple walking events identical in reality, the captured signal variations are not the same in different rounds. To deal with the above problem, RF-Identity develops a hybrid deep learning neural network to extract walking pattern features and body shape features for identification. The workflows are shown in Fig. 8.

4.2.1 Walking Pattern Feature Extraction. To extract the walking pattern features, the first problem encountered is how to effectively incorporate the readings from spatially deployed tags. To tackle this problem, we propose a weight-based multi-channel network to automatically assign weights for each tag and then incorporate readings from them to extract the walking pattern features. Specifically, we first design an attention block that can automatically learn to focus more attention on those tags whose readings are more sensitive to user's walking. Let X_p and X_r respectively denote the phase matrix and amplitude matrix with a size of $K \times M$, where K is the number of tags and M is the length of the phase or amplitude reading stream.[1] We take the phase stream as an example and the weight w_p^k for the k-th tag is computed by:

$$w_p^k = \frac{\exp\left(e_p^k\right)}{\sum_{j=1}^{K} \exp\left(e_p^j\right)} \tag{4}$$

and e_p^k is defined as:

$$e_p^k = f\left(g_p^k, X_p^k\right) \tag{5}$$

[1]We perform a data interpolation operation to make the lengths of phase/amplitude readings from different tags the same.

where f is the three-fully-connected layers, and g_p^k is the weight vector of the hidden layers. We then perform a matrix multiplication between X_p and W_p to obtain the weighted phase matrix Z_p as follows:

$$Z_p = X_p \otimes W_p \tag{6}$$

where $W_p = \left(w_p^1, w_p^2 \cdots, w_p^K\right)^T$, and T is the transpose operation. In the same way, we can obtain the weighted amplitude matrix Z_a. With this attention mechanism, RF-Identity can speed up the network to extract features effectively. To fully combine the phase and amplitude information, we perform a concatenation operation as follows:

$$C = Z_p \oplus Z_a \tag{7}$$

We then put C into a ResNet module to acquire distinctive features from phase and amplitude information. The reason for selecting the residual block is that it comprises the identity shortcut connections, which keeps all information passing through freely in Fig. 8. Specifically, the shortcut connections can be written as:

$$V = H\left(C, w_C\right) + C \tag{8}$$

where $H\left(C, w_C\right)$ is a function to learn the residual mapping, w_C denotes the parameter of each layer, C is the input vector and and V is the output vector. Then an average pooling layer and a flatten layer are employed to output the extracted walking pattern features $V_{walking}$. To visualize the learned features, we use t-SNE [6] to project them into a two-dimensional feature space. Fig. 9 depicts the features of 15 walking human subjects. It shows that different subjects have different feature distributions, which can be used to identify subjects. However, the feature distributions of some users have overlapped regions, leading to user identification errors. We need to derive more features to improve the accuracy of person identification.

4.2.2 Body Shape Feature Extraction. We design a body shape estimation module and combine the extracted features togetehr with the walking pattern features for user identification. The basic idea is to utilize amplitude attenuation features collected from spatially deployed RFID tags. Instead of using the absolute amplitude values X_{A_n} which can be affected by hardware noise, we use the amplitude change $\tilde{X}_{A_n}$ before and after the human subject appears to remove the common hardware noise:

$$\tilde{X}_{A_n}^k = \left| X_{A_n}^k - \bar{X}_{A_{free}}^k \right| \tag{9}$$

where n is the sample index, k is the tag index, and $\bar{X}_{A_{free}}^k$ denotes the average amplitude reading of the k-th tag when the sensing area is empty.

Next, we detail how to capture body shape information based on amplitude change $\tilde{X}_{A_n}$. Inspired by previous work [44] which uses amplitude measurements to perform radio tomographic imaging (RTI) for localization, RF-Identity employs amplitude change-based RTI method to obtain the body shape information. Note that different from the previous work [44] which deploys many transceivers at each side of the area to generate multiple pairs of transmitter-receiver around the target, RF-Identity only requires an RFID reader antenna at one side and a few tags at the other side. It exploits only a few RF links to obtain the body reflected snapshot and combine multiple snapshots across time to recover the body shape tomographic image. We collect amplitude change stream $\tilde{X}_A^k$ for each tag and form the amplitude change matrix $\tilde{X}_A$ as below:

$$\tilde{X}_A = \begin{bmatrix} \tilde{X}_{A_1}^1 & \tilde{X}_{A_2}^1 & \cdots & \tilde{X}_{A_M}^1 \\ \tilde{X}_{A_1}^2 & \tilde{X}_{A_2}^2 & \cdots & \tilde{X}_{A_M}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{X}_{A_1}^K & \tilde{X}_{A_2}^K & \cdots & \tilde{X}_{A_M}^K \end{bmatrix} \tag{10}$$

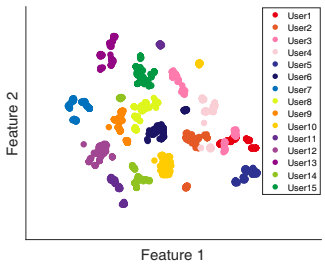


Fig. 9. The t-SNE visualization of walking pattern features for 15 walking subjects.

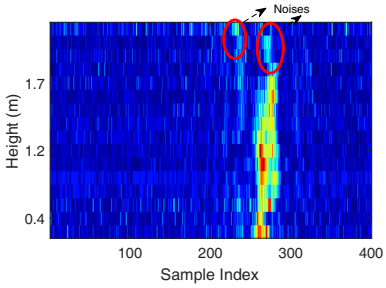


Fig. 10. The constructed tomographic image (i.e., the body shape feature) for a walking subject.

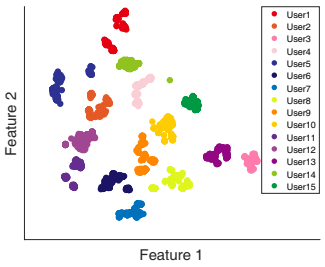


Fig. 11. The t-SNE visualization of combined features for 15 walking subjects.

Then, based on a linear relationship between the amplitude change and the pixel value, we have:

$$\tilde{X}_A = WF + N \tag{11}$$

where W is a transformed weight matrix with a dimension of $K \times M$. The weight of each pixel is generally determined by the ellipsoid-based method [23]. F is the objective function that denotes the pixel value matrix of the tomographic image and N is the noise vector. To obtain F, we transform Equation (11) into a least-squared-error optimization problem:

$$\arg\min_{F} \left\| WF - \tilde{X}_A \right\|_2^2 \tag{12}$$

Equation (12) is an ill-posed problem, which means a small amount of noise in $\tilde{A}$ is magnified [23]. To address this problem, we add a regularization term and rewrite it as follows:

$$\left\| WF - \tilde{X}_A \right\|^2 - \beta \left\| IF \right\|^2 \tag{13}$$

where I is an identity matrix and β is a regularization parameter. We can now obtain the explicit solution $\hat{F}$ based on [43] as below:

$$\hat{F} = \left(W^T W + \beta I \right)^{-1} W^T \tilde{X}_A \tag{14}$$

In this paper, the regularization parameter β is an empirical value of 55.

Fig. 10 shows the constructed tomographic image when a subject with a 1.68 m height walks through the sensing area. We observe that the constructed image is able to capture the body shape information. The warmer the pixel color, the larger the value is. The intuition is that when human body blocks the radio links, it has different impacts on different tags. To further remove the scattered points in the red circles in Fig. 10 to obtain more accurate body shape information, we leverage an encoder block to process the constructed tomographic image as follows:

$$Z_F = Encoder\left(\hat{F}, w_F \right) \tag{15}$$

where *Encoder* consists of three convolution layers and w_F denotes the parameters to be learned. Then, Z_F is put into an average pooling layer and a flatten layer to extract the learned body shape information Z_{body}.

4.2.3 Feature Fusion. Based on the extracted walking pattern features $V_{walking}$ and body shape features Z_{body}, we perform feature fusion by concatenating them as follows:

$$U = V_{walking} \oplus Z_{body} \tag{16}$$

To predict the labels of test subjects, we use two fully connected layers to map the feature representation U into a new space S. Then, we use a softmax layer to calculate the probability $\hat{y}$. Finally, RF-Identity uses a loss function module to compute the similarity between the predicted label $\hat{y}_i$ and the ground-truth y_i. Here, we use a cross-entropy to calculate the loss L_1 as:

$$L_1 = -\frac{1}{m_b}\sum_{i=1}^{m_b} y_i \log\left(\hat{y}_i\right) + \kappa \mathrm{O} \tag{17}$$

where m_b is the size of the batch, O is a regularizer to avoid overfitting and κ is the hyper-parameter.

We train the proposed model on the training set by iteratively using the backward propagation to minimize the cross-entropy. By applying the proposed method, we visualize the learned features for 15 subjects with t-SNE in Fig. 11. We can observe that different subjects have more distinct feature distributions compared with those in Fig. 9. This result demonstrates that the hybrid deep learning model can effectively acquire features for classification by combining the walking patterns and body shape information.

4.3 Data Augmentation

In this part, we present how to generate synthesized training samples to reduce the data collection cost. We design a data augmentation scheme that includes three signal transformations of the original data. That is to say, when we input a data X, the outputs are three variants of X. Specifically, the three signal transformations are as follows:

- **Jitter:** Jitter is a method to add random noise to the original data without changing the length of X. It can imitate the noise caused by environment, software and hardware. Thus, the synthesized jitter data can enhance the capability of the model against different categories of noise.
- **Time-Warped:** Time-Warped transformation changes the temporal scale of the received data by varying the time intervals of the original data samples. The received signal stream may get stretched or compressed to represent different walking velocities.
- **Different numbers of readings for each tag:** When there are multiple tags, the numbers of readings collected from each tag are not the same in a same interval. Therefore, to imitate this phenomenon, we also randomly generate different numbers of readings for each tag. After that, we perform interpolation to make the the number of readings the same for all the tags.

Based on the proposed data augmentation scheme, we achieve the following benefits: 1) we significantly reduce the amount of human effort in collecting training data; 2) The synthesized data generated by three transformations greatly improves the model's robustness.

4.4 Accommodating New Users

To reduce the cost of re-collecting data for new users, we employ a transfer learning method to speed up the convergence of the model parameters. The observation is that some of the previously trained parameters of the model are still valid for the new subjects. So we can transfer learned knowledge from previously trained users (source domain) to new users (target domain).

Fig. 12 illustrates how this process works in detail. The transfer learning scheme includes two parts: the frozen module, and the updating module. The frozen module is to fix the pre-trained parameters of the hybrid deep learning model, thus substantially reducing the time and human labor cost for data collection. The updating module updates the model weights quickly by learning the information from the target domain with two fully-connected layers.

Note that the dimension of the updating module's last layer relates to the number of predicted labels. Suppose the parameters are trained for n subjects in advance, once a new subject joins in, we need to change the last

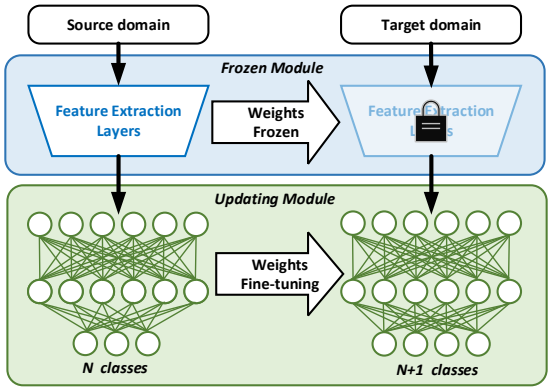
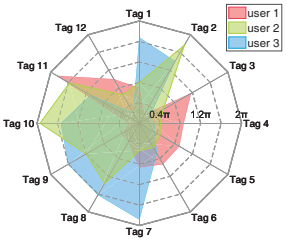
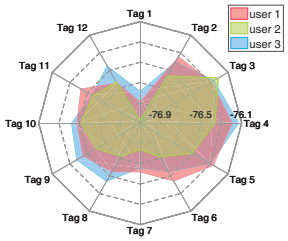


Fig. 12. The transfer learning model for new users.

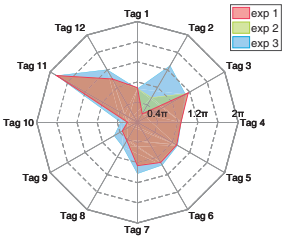


(a) Phase

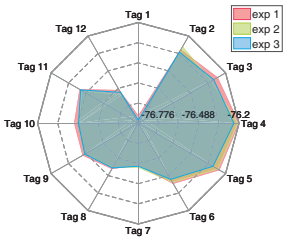


(b) Amplitude

Fig. 13. The phase and amplitude profiles of different subjects.



(a) Phase



(b) Amplitude

Fig. 14. The phase and amplitude profiles of one subject for three repeated experiments.

layer's dimension from n to $n + 1$ in the training phase and randomly initialize the parameters. Based on the above steps, we can reduce the training effort for new users and still retain a good performance.

4.5 Static User Feature Extraction

In this subsection, we introduce how to extract features from a static user. The basic idea is to use spatially deployed tags to capture the distinct effects on wireless signals induced by users of different biometrics (e.g., height). We conduct a benchmark experiment with three subjects standing still at a fixed position.[2] The phase and amplitude readings are plotted in Fig. 13. We can observe that different users exhibit different phase and amplitude profiles constructed with data collected from spatially deployed tags. Moreover, we let one of the subjects stand still at a fixed position, and repeatedly collect readings three times within a time interval of five minutes. The phase and amplitude profiles are shown in Fig. 14. We can see that the readings of some tags are stable while others are not.

To overcome the above issue, we introduce a weight-based distance method to help differentiate profiles of different users. Specifically, we introduce two weights α and η to estimate the profile distance for each subject, where α represents the ability to identify different users, and η stands for the similarity between the training and

[2]In the sensing area, 12 tags and a directional antenna are placed at two sides of the door. The detailed deployment is shown in Fig. 16.

Table 1. Eight statistical features

ID	Statistical feature	ID	Statistical feature
1	Standard deviation	2	Mean absolute deviation
3	Median absolute deviation	4	Interquartile range
5	Root mean square	6	Entropy
7	Skewness	8	Kurtosis

testing profiles. The two weights are respectively described as follows:

$$\alpha_k = \frac{\sum_{i}\sum_{h\neq b} dist\left(P_{k,S}^{i,h}, P_{k,S}^{i,b}\right)}{\sum_{i\neq j}\sum_{h,b} dist\left(P_{k,S}^{i,h}, P_{k,S}^{j,b}\right)} \tag{18}$$

$$\eta_k = \frac{\text{var}\left(dist\left(P_{k,T}, P_{k,S}^{i,h}\right)\right)}{\sum_{k=1}^{K}\text{var}\left(dist\left(P_{k,T}, P_{k,S}^{i,h}\right)\right)} \tag{19}$$

where $dist\,(,)$ is the Euclidean distance, $P_{k,S}^{i,h}$ is the phase value of the $k\text{-}th$ tag in the $h\text{-}th$ recorded profile for person i and $P_{k,T}$ is the $k\text{-}th$ phase value of the testing profile. When α and η are smaller values, the $k\text{-}th$ tag has a stronger distinguishing capability and is more robust. Thus, we formulate the phase profile distance as follows:

$$ED = \sqrt{\sum_{k=1}^{K}\alpha_k \cdot \eta_k \cdot dist\left(P_{k,T}, P_{k,S}\right)} \tag{20}$$

Similarly, we can obtain the amplitude profile distance. Moreover, RF-Identity calculates another eight statistic features (listed in Table 1) of phase and amplitude profiles, respectively. Then we incorporate them with the profile distance features as the final static user features. Fig. 15 illustrates the merged features of 15 subjects, which have excellent discrimination capability.

4.6 Person Identification and Spoof Detection

In this section, we introduce how to perform person identification and spoof detection. Person identification is to identify users registered in the training dataset, and spoof detection refers to detecting illegal users. Specifically, RF-Identity feeds the extracted dynamic/static user features into a SVM [31] with a radial basis function (RBF) kernel function to identify dynamic/static users. To perform spoof detection, RF-Identity leverages SVDD algorithm [26] to find a minimum hypersphere. The hypersphere contains pre-collected legal users' data and excludes illegal users' data. When a test user is in the hypersphere, RF-Identity determines that the user is legitimate and illegal otherwise.

5 IMPLEMENTATION

Hardware implementation: The system setup is shown in Fig. 16. It contains an Impinj Speedway R420 reader with one directional antenna (9 dBi gain; 70° elevation and azimuth beam width). The reader operates at a frequency of 924.375 MHz and 12 ALN-9640 tags are employed to form a tag array. The reader has a sampling rate of 140 samples per second. The horizontal distance between the antenna and the middle point of the tag array is 1.5 m. The height of the antenna is 1.2 m above the ground. Note that the spacings between adjacent tags

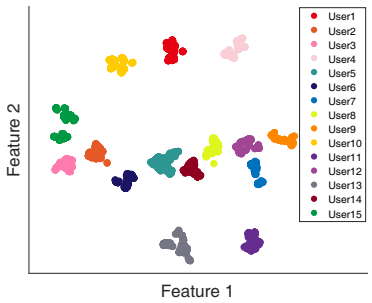


Fig. 15. The t-SNE visualization of learned features for 15 static subjects.

Fig. 16. Experimental scenario.

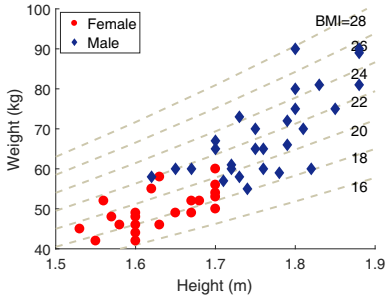


Fig. 17. The distribution of volunteers.

are not uniform. The first four tags are deployed at a height ranging from 0.3 m to 1.5 m at a step size of 0.4 m. The rest eight tags are placed at a height ranging from 1.55 m to 1.9 m at a step size of 0.05 m.

Backend implementation: The R420 reader collection program is implemented with *C#* code on a Lenovo Thinkpad laptop with an Intel i7-6700HQ CPU and 32GB memory. The software is compatible with the Low-level Reader Protocol (LLRP) Toolkit to output phase and amplitude readings. The proposed deep learning algorithm is implemented in Python, based on Keras framework with Tensorflow backend. It is trained at a remote server with an Intel Xeon(R) CPU E5-2620 v3, a 64GB memory and a NVIDIA TITAN RTX GPU.

Data collection: In our default experiments, we recruit 50 volunteers (28 males and 22 females). As shown in Fig. 17, different subjects have different heights and weights. To perform person identification, we conduct two rounds of experiments. In the first round, we ask each volunteer to walk across the sensing area[3] for 80 times, and collect RFID measurements throughout the process. In the second round, each volunteer stands at a fixed location 10 cm before the LoS for a period of 5 seconds for 50 times and we collect the RFID measurements. We refer to the above collected dynamic and static data as dataset 1. We use a table to record the identity of the user for each experiment and label all the data manually. For all the experiments, we ensure that only one volunteer is in the sensing area and no other person moves around.

6 PERFORMANCE

6.1 Overall Performance

6.1.1 Person identification accuracy. To evaluate the performance of dynamic person identification, we employ 10-fold cross-validation to split collected dynamic samples of dataset 1. In other words, in each fold, we select $1/10 \times 80 = 8$ samples of each user for testing and the remaining $(1 - 1/10) \times 80 = 72$ samples for training. Fig. 18 presents the confusion matrix for 50 subjects. RF-Identity achieves an average accuracy of 94.2% in dynamic person identification. Some subjects can be identified at 100% accuracy, which demonstrates the effectiveness of the proposed method. To evaluate the performance of identifying static users, we use 80% of the static measurements of dataset 1 for training and 20% for testing. The results are shown in Fig. 19 and RF-Identity performs well for static user identification with an average accuracy of 95.9%.

6.1.2 Spoof detection accuracy. We now evaluate the spoof detection performance. Specifically, we choose h subjects as legitimate users and the remaining $50 - h$ subjects as illegal spoofers, where h varies from 5 to 45 at a step size of 5. RF-Identity employs the SVDD classifier to detect spoofers. The results are presented in Fig. 20. We can see that the spoof detection accuracy goes up as the number of legitimate users increases in both cases.

[3]The default experiment scenario is at the entrance of the lab.

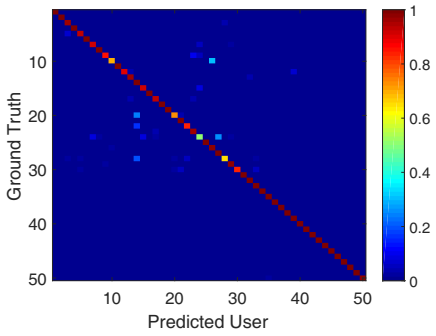


Fig. 18. The confusion matrix in dynamic case.

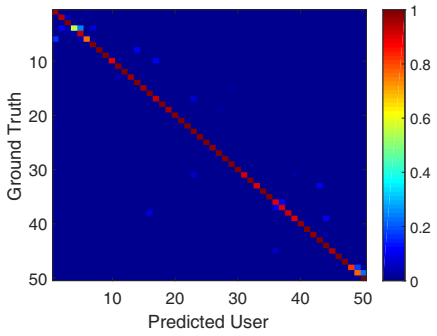


Fig. 19. The confusion matrix in static case.

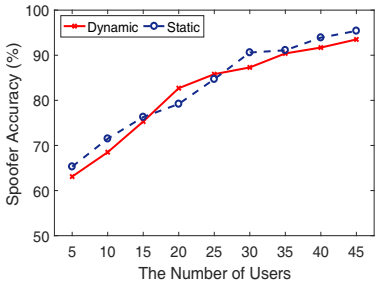


Fig. 20. The performance of spoof detection.

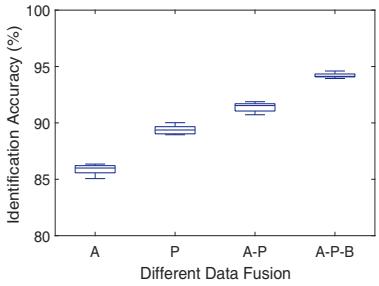


Fig. 21. The performance of feature fusion in dynamic case.

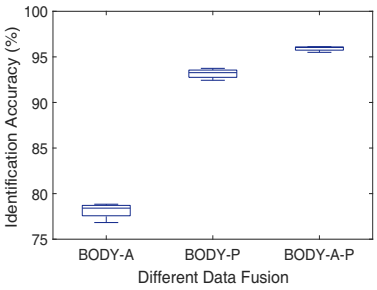


Fig. 22. The performance of feature fusion in static case.

From the results, we can conclude that RF-Identity can effectively recognize illegal users as long as the size of the training data is not too small.

6.1.3 Verification of the feature fusion. To assess the performance of the feature extraction model for dynamic and static users, we conduct a set of benchmark experiments by changing the model inputs with dataset 1. Specifically, in the dynamic case, we use four different inputs: (1) phase-based walking pattern feature (P); (2) amplitude-based walking pattern feature (A); (3) the combination of (1) and (2) (A-P); and (4) the combination of (3) and body shape feature (A-P-B). Fig. 21 plots the identification accuracy under different inputs. The results depict that phase-based performance is better than amplitude-based, which implies that phase readings contain more fine-grained information. And the fusion of phase and amplitude achieves a better performance. By further including the body shape features, the A-P-B combination achieves the best performance. Thus, the proposed hybrid deep learning model can efficiently combine the walking pattern and body shape features to achieve a better performance. In the static user case, we obtain similar results in Fig. 22. The combination of phase and amplitude features can achieve a higher identification accuracy.

6.1.4 Verification of the data augmentation method. We analyze the person identification performance with and without the data augmentation (DA) method. Specifically, we change the number of dynamic training samples from 3 to 24 for each subject. Fig. 23 depicts the average identification accuracy. The accuracy improves when the number of training sample increases. It increases from 49.5% to 79.8% as the number of training sample increases from 3 to 24 without augmentation (w/o DA), and from 68.2% to 94.1% with augmentation (w/ DA). For the same

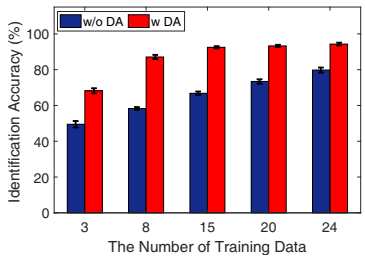
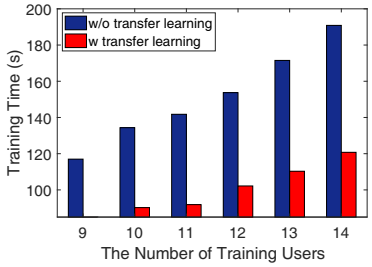
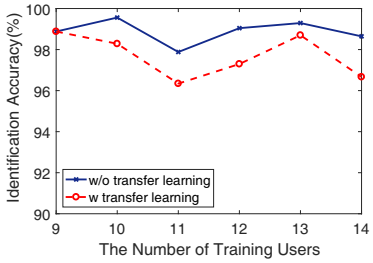


Fig. 23. The performance of data augmentation (DA) method.



(a) The training time of hybrid deep learning module.



(b) The accuracy of hybrid deep learning module.

Fig. 24. The performance of transfer learning scheme.

number of training sample, the accuracy increases a lot with augmentation (w/ DA) when compared to without augmentation (w/o DA). For instance, with 15 training samples, RF-Identity can achieve an average accuracy of 92.5% w DA and 66.8% w/o DA. In conclusion, the data augmentation scheme can effectively improve system performance by generating more training data. Data augmentation also reduces the amount of data collection effort. To balance the data collection overhead and accuracy, we suggest to collect 15 training samples to achieve a reasonably good performance.

6.1.5 Verification of the transfer learning method. To verify the effectiveness of the transfer learning method, we initialize the hybrid deep learning model by training data of 9 subjects and freeze the trained parameters. Then we apply data collected from various numbers of training subjects (9 to 14) to evaluate the system performance with and without applying the transfer learning method. The results are shown in Fig. 24. Fig. 24(a) presents the training time of the hybrid deep learning model. It manifests that the transfer learning method remarkably saves the training time. Fig. 24(b) exhibits the average identification accuracy for different numbers of users. The accuracy with transfer learning is always above 96% and is comparable to that achieved without transfer learning. Thus, the transfer learning method can considerably reduce the training cost and retain a higher identification accuracy.

6.2 Diverse factors on RF-Identity

In this subsection, we investigate the factors affecting the system performances.

6.2.1 Effect of the number of users to be differentiated. We change the number of tested users from 10 to 50 to evaluate the impact on identification accuracy. The performance is shown in Fig. 25. We can see that the identification accuracy decreases with more users to be differentiated. In particular, when there are only 10 targets to be differentiated, RF-Identity achieves an identification accuracy of 98.7% and 99.6% for dynamic and static cases, respectively. When there are 50 targets, the accuracy slightly decreases to 94.2% and 95.9% respectively for dynamic and static cases. This result is expected because with more targets to be differentiated, the average difference between targets' features gets smaller.

6.2.2 Effect of different clothes. In practical scenarios, the same subject may wear different clothes in different days. To explore the effects of different clothes, we ask 10 volunteers to perform three rounds of experiments wearing three different clothes. In each round, we ask the volunteers to wear the same clothe (same brand and model of different sizes), and walk naturally for 30 times and stand for a period of 5 seconds for 30 times. Note that these clothes are different from those used in the training phase. Then, we use these collected samples to evaluate the model trained by dataset 1. Fig. 26 shows the identification performance. The accuracy slightly

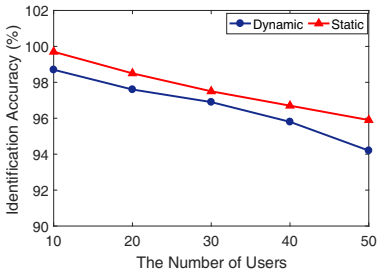


Fig. 25. The performance with varying number of users.

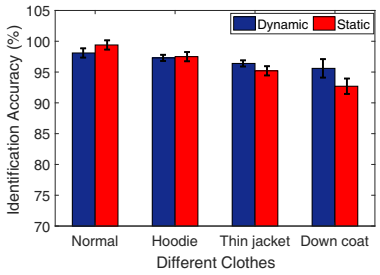


Fig. 26. The performance with different clothes.

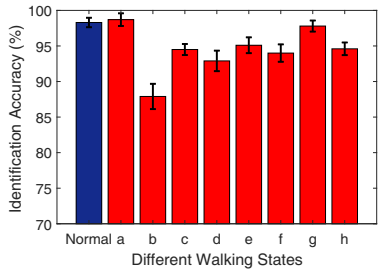


Fig. 27. The performance with different walking events.

Table 2. Different walking events

ID	Walking event	ID	Walking event
a	slow walking	b	fast walking
c	trajectory 1	d	trajectory 2
e	make phone call	f	drink water
g	carry a laptop	h	carry a backpack

decreases from 98.3% to 95.6% in the dynamic case and 99.4% to 92.7% in the static case. The result suggests that the system is more robust in the dynamic case than the static case when the subject wears different clothes. We believe the reason is that the dynamic user feature mainly depends on the walking pattern. In contrast, the static user feature is more related to the RF attenuation induced by the body, which is more affected by the clothe. Overall, RF-Identity still achieves a high accuracy over 92% with different clothes in both dynamic and static cases.

6.2.3 Effect of different walking events. We evaluate the effects of different walking events by letting the same 10 subjects walk through the sensing area at different velocities, following different trajectories and carrying different objects. For each event, the subject is asked to walk through the entrance for 30 times. Table 2 describes the detailed walking events. As shown in Fig. 27, except for the fast walking which achieves a decreased accuracy of 87.9%, RF-Identity achieves an average accuracy of 95.7% in other events. These results indicate that the proposed model can effectively extract the dynamic user features and is robust against varying walking events.

6.2.4 Effect of different environments. We estimate the system performance when the system is deployed in two new environments: a corridor entrance and an office entrance. Note that the experiment setups are kept the same as that in the default lab environment. In the new environment, we ask each volunteer to walk naturally across the sensing area for 30 times and stand still for a period of 5 seconds for 30 times. Then, we evaluate the identification accuracy using the model trained with dataset 1. Fig. 28 shows the identification performance in the two new environments. We also include the identification performance in the default lab environment for comparison. The performance in the corridor and office environments decreases compared to performance in the default lab environment. This is because the extracted features are environment-dependent. Therefore, using the model trained in one environment to work with data collected in another environment can cause performance degradation.

6.2.5 Effect of the number of tags. We now explore the impact of tag number on RF-Identity's performance. For each tag deployment, we employ 15 volunteers to walk naturally across the sensing area for 80 times and

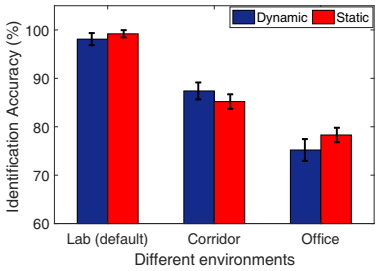


Fig. 28. The performance with different environments.

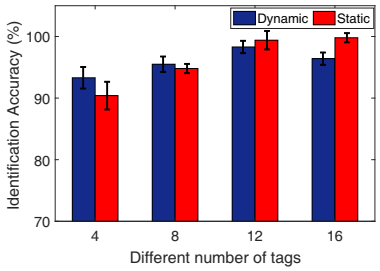


Fig. 29. The performance with different number of tags.

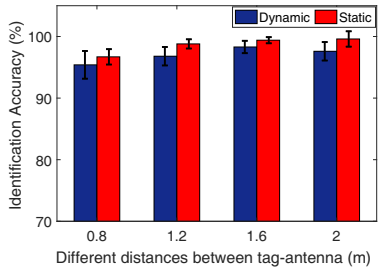


Fig. 30. The performance with different horizontal distances between the reader antenna and tag-array.

Table 3. Comparison with different user identification systems.

System	Signal	Users	Accuracy
WiFi-ID	Wi-Fi	2-6	93%-77%
Wiwho	Wi-Fi	2-6	92%-80%
WiFiU	Wi-Fi	50	79.28%
RF-Identity	RFID	50	94.21%

stand still for 50 times. Then, we use the data collected to re-train the model and evaluate the performance with the re-trained model. Fig. 29 shows that the identification accuracy increases with the increase of tag number. The reason is that more tags can provide richer spatial information of a user. Thus the system can obtain more fine-grained behavioral and physical features for person identification. In this paper, we adopt 12 tags based on our empirical study.

6.2.6 Effect of the horizontal distance between the reader antenna and tag array. In practice, the width of an entrance varies significantly[1]. Because we deploy the RFID reader at one side of the entrance and the tags at the other side, the entrance width determines the distance between the RFID reader and tags. We evaluate the impact of the entrance width on RF-Identity's performance. We vary the horizontal distance between the reader and tags from 0.8 m to 2 m at a step size of 0.4 m. We ask 15 volunteers to walk naturally across the sensing area for 80 times and stand still for 50 times. Then, we use the collected data to re-train the model and evaluate the performance with the re-trained model. Fig. 30 shows that the reader-tag distance change does not affect the accuracy much. RF-Identity achieves a slightly higher accuracy when the reader-tag distance is larger. We believe this is because the static signal strength decreases more than the dynamic signal when the tag-reader distance increases [48]. The dynamic signal is usually much weaker than the static signal due to reflection. Thus, when the two signals both decrease but the static signal decreases more, the signal strength difference between the two signals becomes smaller. Therefore, the variation of the dynamic signal can now cause a larger variation on the composite signal.

6.2.7 Comparison with prior works. We compare RF-Identity with three state-of-the-art Wi-Fi-based user identification systems, Wiwho[52], WiFiU[40] and WiFi-ID[54] in Table 3. We can see that RF-Identity outperforms these approaches in terms of accuracy. Furthermore, the existing systems cannot identify static human targets. In addition, we implement one system above (WiFiU) using one Wi-Fi transceiver pair. We use a TP-link WDR7500 Wi-Fi router that supports IEEE 802.11n protocol as the transmitter and a mini-PC with an Intel 5300 wireless card

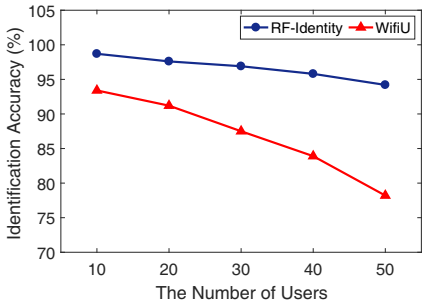


Fig. 31. Comparison to prior work.

as the receiver to collect Wi-Fi CSI readings. The wireless router works at 5 GHz band with a channel bandwidth of 20 MHz. The transmitter and receiver are placed with a distance of 1.5 m in-between at the same height of 1.2 m, which is the same as the RFID deployment. The packet transmission rate is 1000 packets per second. We ask 50 subjects to naturally walk through the sensing area for 30 times. The comparison results are shown in Fig. 31. We can see that RF-Identity outperforms WiFiU and the amount of improvement increases with a larger number of tested users.

7 DISCUSSION

We discuss the limitations and some interesting observations in this section.

1) *Multi-user scenarios.* Multi-user sensing is a well-known challenge in RF sensing because the reflection signals from multiple targets get mixed at the receiver, interfering with each other. The current version of RF-Identity works well with a single human target. When there are multiple targets, the system still works if the targets are far away from each other. If multiple targets are close to each other, it is challenging for our system to work. We believe some recent progress in multi-dimensional signal processing [47] can help to achieve multi-user sensing. It is a challenging yet interesting research direction to deal with multiple targets in our future work.

2)*Environment diversity.* To achieve highly accurate person identification, RF-Identity requires the test environment to be the same as training environment because the model trained is environment-dependent. Besides, RF-Identity requires the user to stay at a fixed location for static identification since the phase and amplitude profiles will change when the user location changes. We believe more advanced deep learning methods such as transfer learning and adversarial learning [3, 53] can be employed to mitigate these limitations.

3) *Tag deployment.* To balance the system's accuracy and latency, in RF-Identity, the tags are deployed non-uniformly in the vertical direction. The reason is that the tag reading rate decreases when the number of deployed tag increases. Therefore, we deploy more tags on the top to obtain fine-grained human body shape information (e.g., height), and deploy less tags at the bottom to extract the gait information. For future design, we believe 2-dimensional tag array can capture even richer information of the target. For example, the current linear tag array enables us to obtain the height information of the target. With a second dimension of tags, we are able to obtain the thickness of the target. Another interesting direction worth exploring is employing the coupling effect among tags to further enhance the sensing performance. One recent work [46] successfully employed this "bad" coupling effect to improve the sensing performance. We discovered in the paper that with more tags deployed, we can obtain more accurate human body shape information. However, we observe that when we deploy more tags, the spacing between tags becomes smaller and eventually coupling effect is induced.

8 RELATED WORK

In this section, we review the related literature in person identification and other RFID-based sensing work.

8.1 Camera-based person identification

Camera-based systems [25, 59, 60] are widely used in person identification. Although a high accuracy can be achieved, they require good lighting conditions that can raise privacy concerns. Besides, these methods can be vulnerable to replay attacks [33]. In contrast, RF-Identity is independent of lighting conditions and does not raise privacy concerns.

8.2 Sensor based person identification

Many systems utilize sensors, such as fingerprint sensor, vibration sensor, and microphone sensor for person identification. A recent work [10] collects users' fingerprints to verify identity but can be fooled by a thin film. Zhu *et al.* [50] leverages iris scan sensors to authenticate user identity but the hardware cost is high. FootprintID [22] performs robust indoor pedestrian identification by using vibration sensors to detect vibrations induced by footsteps. However, it requires many sensors to cover the sensing area. BreathLive [12] performs person authentication using microphone sensors to find the correlation between sounds and chest motion, which requires users to carry devices. Unlike the above systems, RF-Identity is a non-intrusive system which does not require users to carry any device and is implemented on relatively cheap commercial hardware.

8.3 RF-based person identification

Various RF technologies including Radar [27, 34], FMCW [2, 57, 58], Wi-Fi [16, 20, 28, 30, 53] and RFID [11, 56] are utilized to perform person identification. FormaTrack[15] exploits a radar to capture body shape as a biometric for person identification. RF-Capture[2] uses a dedicated FMCW to capture the human figure behind the wall. These systems achieve accurate identification, yet still require dedicated and expensive devices. Gait-based approaches [40, 52, 54] require the subject to keep the same walking pattern for high accuracy, and the performance decreases as the number of test users increases. RF-Mehndi [56] leverages the tag coupling effect for person identification. It requires the user to carry a dedicated card with a well-designed tag array printed on it. Au-Id [11] proposes a deep learning model to extract knocking features, which requires a massive amount of training data. Some previous works [9, 13, 36] can identify static users but require a large bandwidth to obtain richer features. Unlike existing schemes, RF-Identity uses commodity hardware without a need of large bandwidth. RF-Identity is able to identify both dynamic and static persons in a contact-free manner.

8.4 RFID-based sensing

RFID technology has been widely applied in many sensing applications, such as localization [37, 39], activity and gesture recognition [5, 29, 35], object interaction detection [17, 55] and target material identification [38]. FaHo [49] uses radio frequency holograms to locate RFID tags. Grfid [61] performs accurate and robust gesture recognition by developing a weighted DTW method. EUIGR[51] employs adversarial learning to achieve real-time gesture recognition. Rio [24] and Tagtag [45] utilize tag coupling effect to sense touch gesture and target material, respectively. Different from the state-of-the-arts, RF-Identity fuses phase and amplitude information to perform dynamic/static person identification.

9 CONCLUSION

This paper introduces a person identification system RF-Identity which is capable of identifying not just dynamic users but also static users using commodity RFID hardware. In the dynamic case, RF-Identity proposes a hybrid deep learning framework to integrate the behavioral features and physiological features for person identification.

In the static case, RF-Identity utilizes tag diversity in spatial domain to construct phase and amplitude profiles to distinguish static users. Extensive experiments demonstrate RF-Identity can achieve accurate and robust person identifications in a non-intrusive way.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61802308 and 61772422. This work is also partially supported by NSFC A3 Foresight Program Grant 62061146001, the Shaanxi International Science and Technology Cooperation Program under grant agreement 2019KW-004, 2020KW-008, and 2019KWZ-05, the ShaanXi Science and Technology Innovation Team Support Project under grant agreement 2018TD-026.

REFERENCES

[1] [n. d.]. What Is The Standard Size Of Doors And Windows In Residential Buildings? https://www.acivilengineer.com/standard-size-of-doors-and-windows/.

[2] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. 2015. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 219. doi:10.1145/2816795.2818072.

[3] Xi Chen, Hang Li, Chenyi Zhou, Xue Liu, Di Wu, and Gregory Dudek. 2020. FiDo: Ubiquitous Fine-Grained WiFi-based Localization for Unlabelled Users via Domain Adaptation. In *Proceedings of The Web Conference 2020*. 23–33. doi:10.1145/3366423.3380091.

[4] Jan K Chorowski, Dzmitry Bahdanau, Dmitriy Serdyuk, Kyunghyun Cho, and Yoshua Bengio. 2015. Attention-based models for speech recognition. In *Advances in neural information processing systems*. 577–585. doi:10.1016/0167-739X(94)90007-8.

[5] Wang Chuyu, Liu Jian, Chen Yingying, Xie Lei, Liu Hongbo, and Lu Sanglu. 2018. RF-Kinect: A Wearable RFID-based Approach Towards 3D Body Movement Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1 (June 2018). doi:10.1145/3191773.

[6] Laurens Van Der Maaten and Geoffrey E Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9 (2008), 2579–2605. doi:10.1007/s10846-008-9235-4.

[7] Chao Feng, Jie Xiong, Liqiong Chang, Ju Wang, Xiaojiang Chen, Dingyi Fang, and Zhanyong Tang. 2019. WiMi: Target Material Identification with Commodity Wi-Fi Devices. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 700–710. doi:10.1109/ICDCS.2019.00075.

[8] Huan Feng, Kassem Fawaz, and Kang G Shin. 2017. Continuous Authentication for Voice Assistants. (2017), 343–355. doi:10.1145/3117811.3117823.

[9] Tianbo Gu, Zheng Fang, Zhicheng Yang, Pengfei Hu, and Prasant Mohapatra. 2019. mmSense: Multi-Person Detection and Identification via mmWave Sensing. In *Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems*. 45–50. doi:10.1145/3349624.3356765.

[10] Ke Han, Zizheng Wang, and Zilong Chen. 2018. Fingerprint Image Enhancement Method based on Adaptive Median Filter. In *2018 24th Asia-Pacific Conference on Communications (APCC)*. doi:10.1109/APCC.2018.8633498.

[11] Anna Huang, Dong Wang, Run Zhao, and Qian Zhang. 2019. Au-Id: Automatic User Identification and Authentication through the Motions Captured from Sequential Human Activities Using RFID. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 48. doi:10.1145/3328919.

[12] Chenyu Huang, Huangxun Chen, Lin Yang, and Qian Zhang. 2018. Breathlive: Liveness detection for heart sound authentication with deep breathing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 12. doi:10.1145/3191744.

[13] Shekh Md Mahmudul Islam, Olga Borić-Lubecke, Yao Zheng, and Victor M Lubecke. 2020. Radar-Based Non-Contact Continuous Identity Authentication. *Remote Sensing* 12, 14 (2020), 2279. doi:10.3390/rs12142279.

[14] Wang Ju, Xiong Jie, Hongbo Jiang, Xiaojiang Chen, and Dingyi Fang. 2017. D-Watch: Embracing "Bad" Multipaths for Device-Free Localization With COTS RFID Devices. *IEEE/ACM Transactions on Networking* PP, 99 (2017), 1–14. doi:10.1145/2999572.2999589.

[15] Avinash Kalyanaraman, Dezhi Hong, Elahe Soltanaghaei, and Kamin Whitehouse. 2017. Forma track: tracking people based on body shape. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 61. doi:10.1145/3130926.

[16] Belal Korany, Chitra R Karanam, Hong Cai, and Yasamin Mostofi. 2019. XModal-ID: Using WiFi for Through-Wall Person Identification from Candidate Video Footage. In *The 25th Annual International Conference on Mobile Computing and Networking*. ACM, 36. doi:10.1145/3300061.3345437.

[17] Hanchuan Li, Can Ye, and Alanson P Sample. 2015. IDSense: A human object interaction detection system based on passive UHF RFID. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2555–2564. doi:10.1145/2702123.2702178.

[18] Hanchuan Li, Peijin Zhang, Samer Al Moubayed, Shwetak N Patel, and Alanson P Sample. 2016. Id-match: A hybrid computer vision and rfid system for recognizing individuals in groups. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.

4933–4944. doi:10.1145/2851581.2889430.
[19] Yan Li, Zilong Wang, Yingjiu Li, Robert Deng, Binbin Chen, Weizhi Meng, and Hui Li. 2019. A Closer Look Tells More: A Facial Distortion Based Liveness Detection for Face Authentication. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security (Asia CCS '19)*. ACM, New York, NY, USA, 241–246. doi:10.1145/3321705.3329850.
[20] Chi Lin, Jiaye Hu, Yu Sun, Fenglong Ma, Lei Wang, and Guowei Wu. 2018. WiAU: An Accurate Device-Free Authentication System with ResNet. In *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9. doi:10.1109/SAHCN.2018.8397108.
[21] Emanuela Marasco and Arun Ross. 2015. A Survey on Antispoofing Schemes for Fingerprint Recognition Systems. *Comput. Surveys* 47, 2 (2015), 28. doi:10.1145/2617756.
[22] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J Mengshoel, Hae Young Noh, and Pei Zhang. 2017. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 89. doi:10.1145/3130954.
[23] Neal Patwari and Piyush Agrawal. 2008. Effects of Correlated Shadowing: Connectivity, Localization, and RF Tomography. In *International Conference on Information Processing in Sensor Networks*. doi:10.1109/IPSN.2008.7.
[24] Swadhin Pradhan, Eugene Chai, Karthikeyan Sundaresan, Lili Qiu, Mohammad Khojastepour, and Sampath Rangarajan. 2017. RIO: A Pervasive RFID-based Touch Gesture Interface. (2017), 261–274. doi:10.1145/3117811.3117818.
[25] Ruijie Quan, Xuanyi Dong, Yu Wu, Linchao Zhu, and Yi Yang. 2019. Auto-ReID: Searching for a Part-aware ConvNet for Person Re-Identification. *arXiv preprint arXiv:1903.09776* (2019). doi:10.1109/ICCV.2019.00385.
[26] Bernhard Scholkopf, John Platt, John Shawetaylor, Alexander J Smola, and Robert C Williamson. 2001. Estimating the Support of a High-Dimensional Distribution. *Neural Computation* 13, 7 (2001), 1443–1471. doi:10.1162/089976601750264965.
[27] Ann-Kathrin Seifert, Moeness Amin, and Abdelhak M Zoubir. 2019. Toward unobtrusive in-home gait analysis based on radar micro-doppler signatures. *IEEE Transactions on Biomedical Engineering* (2019). doi:10.1109/TBME.2019.2893528.
[28] Muhammad Shahzad and Shaohu Zhang. 2018. Augmenting user identification with WiFi based gesture recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–27. doi:10.1145/3264944.
[29] Longfei Shangguan, Zimu Zhou, and Kyle Jamieson. 2017. Enabling gesture-based interactions with objects. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 239–251. doi:10.1145/3081333.3081364.
[30] Cong Shi, Jian Liu, Hongbo Liu, and Yingying Chen. 2017. Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT. In *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 5. doi:10.1145/3084041.3084061.
[31] Johan AK Suykens and Joos Vandewalle. 1999. Least squares support vector machine classifiers. *Neural processing letters* 9, 3 (1999), 293–300. doi:10.1023/A:1018628609742.
[32] A. Tefas, C. Kotropoulos, and I. Pitas. [n. d.]. Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication. *Pattern Analysis Machine Intelligence IEEE Transactions on* 23, 7 ([n. d.]), 0–746. doi:10.1109/34.935847.
[33] Simen Thys, Wiebe Van Ranst, and Toon Goedeme. 2019. Fooling automated surveillance cameras: adversarial patches to attack person detection. *arXiv: Computer Vision and Pattern Recognition* (2019). doi:10.1109/CVPRW.2019.00012.
[34] Fok Hing Chi Tivive, Abdesselam Bouzerdoum, and Moeness G Amin. 2010. A human gait classification method based on radar Doppler spectrograms. *EURASIP Journal on Advances in Signal Processing* 2010 (2010), 10. doi:10.1155/2010/389716.
[35] Chuyu Wang, Lei Xie, Wei Wang, Yingying Chen, Yanling Bu, and Sanglu Lu. 2018. Rf-ecg: Heart rate variability assessment based on cots rfid tag array. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–26. doi:10.1145/3214288.
[36] Fei Wang, Jinsong Han, Feng Lin, and Kui Ren. 2019. WiPIN: Operation-free passive person identification using Wi-Fi signals. In *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 1–6. doi:10.1109/GLOBECOM38437.2019.9014226.
[37] Jue Wang, Deepak Vasisht, and Dina Katabi. 2014. RF-IDraw: Virtual touch screen in the air using RF signals. In *Acm Conference on Sigcomm*. doi:10.1145/2740070.2626330.
[38] Ju Wang, Jie Xiong, Xiaojiang Chen, Hongbo Jiang, Rajesh Krishna Balan, and Dingyi Fang. 2017. TagScan: Simultaneous target imaging and material identification with commodity RFID devices. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. ACM, 288–300. doi:10.1145/3117811.3117830.
[39] Ju Wang, Jie Xiong, Hongbo Jiang, Xiaojiang Chen, and Dingyi Fang. [n. d.]. D-Watch: Embracing Bad multipaths for device-free localization with COTS RFID devices. *IEEE/ACM Transactions on Networking (TON)* 25, 6 ([n. d.]), 3559–3572. doi:10.1109/TNET.2017.2747583.
[40] Wei Wang, Alex X. Liu, and Muhammad Shahzad. 2016. Gait Recognition Using Wifi Signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. New York, NY, USA, 363–373. doi:10.1145/2971648.2971670.
[41] Yuan Wang, Yunhong Wang, and Tieniu Tan. 2004. Combining fingerprint and voiceprint biometrics for identity verification: an experimental comparison. In *International Conference on Biometric Authentication*. Springer, 663–670. doi:10.1007/s10485-013-9317-8.
[42] Yanwen Wang and Yuanqing Zheng. 2018. Modeling RFID Signal Reflection for Contact-free Activity Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4 (Dec. 2018). doi:10.1145/3287071.

[43] Yimin Wei and Michael Ng. 2004. Weighted Tikhonov filter matrices for ill-posed problems. *Applied Mathematics Computation* 149, 2 (2004), 411–422. doi:10.1016/S0096-3003(03)00149-8.

[44] Joey Wilson and Neal Patwari. 2010. Radio Tomographic Imaging with Wireless Networks. *IEEE Transactions on Mobile Computing* 9, 5 (2010), 621–632. doi:10.1109/TMC.2009.174.

[45] Binbin Xie, Jie Xiong, Xiaojiang Chen, Eugene Chai, Liyao Li, Zhanyong Tang, and Dingyi Fang. 2019. Tagtag: material sensing with commodity RFID. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*. ACM, 338–350. doi:10.1145/3356250.3360027.

[46] Binbin Xie, Jie Xiong, Xiaojiang Chen, and Dingyi Fang. 2020. Exploring Commodity RFID for Contactless Sub-millimeter Vibration Sensing. In *the 18th Conference*.

[47] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*. 1–16. doi:10.1145/3300061.3300133.

[48] Tong Xin, Bin Guo, Zhu Wang, Pei Wang, Jacqueline Chi Kei Lam, Victor Li, and Zhiwen Yu. 2018. Freesense: a robust approach for indoor human detection using wi-fi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–23. doi:10.1145/3264953.

[49] Huatao Xu, Dong Wang, Run Zhao, and Qian Zhang. 2019. FaHo: deep learning enhanced holographic localization for RFID tags. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*. ACM, 351–363. doi:10.1145/3356250.3360035.

[50] Zhu Yong, Tieniu Tan, and Yunhong Wang. 2000. Biometric Personal Identification Based on Iris Patterns. *Zidonghua Xuebao/Acta Automatica Sinica* 2, 1 (2000), 2801–2804. doi:10.1109/ICPR.2000.906197.

[51] Yinggang Yu, Dong Wang, Run Zhao, and Qian Zhang. 2019. RFID based real-time recognition of ongoing gesture with adversarial learning. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*. ACM, 298–310. doi:10.1145/3356250.3360045.

[52] Yunze Zeng, Parth H. Pathak, and Prasant Mohapatra. 2016. WiWho: Wifi-based Person Identification in Smart Spaces. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks (IPSN '16)*. Article 4, 12 pages. doi:10.1109/IPSN.2016.7460727.

[53] Jie Zhang, Zhanyong Tang, Meng Li, Dingyi Fang, Petteri Nurmi, and Zheng Wang. 2018. CrossSense: towards cross-site and large-scale WiFi sensing. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM, 305–320. doi:10.1145/3241539.3241570.

[54] Jin Zhang, Bo Wei, Wen Hu, and Salil S Kanhere. 2016. Wifi-id: Human identification using wifi signal. In *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 75–82. doi:10.1109/DCOSS.2016.30.

[55] Tengxiang Zhang, Nicholas Becker, Yuntao Wang, Yuan Zhou, and Yuanchun Shi. 2017. BitID: Easily add battery-free wireless sensors to everyday objects. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 1–8. doi:10.1109/SMARTCOMP.2017.7946990.

[56] Cui Zhao, Zhenjiang Li, Ting Liu, Han Ding, Jinsong Han, Wei Xi, and Ruowei Gui. 2019. RF-Mehndi: A Fingertip Profiled RF Identifier. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 1513–1521. doi:10.1109/INFOCOM.2019.8737419.

[57] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. 2018. Through wall human pose estimation using radio signals. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7356–7365. doi:10.1109/CVPR.2018.00768.

[58] Mingmin Zhao, Yonglong Tian, Hang Zhao, Mohammad Abu Alsheikh, Tianhong Li, Rumen Hristov, Zachary Kabelac, Dina Katabi, and Antonio Torralba. 2018. RF-based 3D skeletons. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. 267–281. doi:10.1145/3230543.3230579.

[59] Zhedong Zheng, Xiaodong Yang, Zhiding Yu, Liang Zheng, Yi Yang, and Jan Kautz. 2019. Joint discriminative and generative learning for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2138–2147. doi:10.1109/CVPR.2019.00224.

[60] Zhen Zhu, Tengteng Huang, Baoguang Shi, Miao Yu, Bofei Wang, and Xiang Bai. 2019. Progressive Pose Attention Transfer for Person Image Generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2347–2356. doi:10.1109/CVPR.2019.00245.

[61] Yongpan Zou, Xiao Jiang, Jinsong Han, Kaishun Wu, Li Yun, and Lionel M. Ni. 2017. GRfid: A Device-Free RFID-Based Gesture Recognition System. *IEEE Transactions on Mobile Computing* 16, 2 (2017), 381–393. doi:10.1109/TMC.2016.2549518.