Smart User Authentication through Actuation of Daily Activities Leveraging WiFi-enabled IoT

Cong Shi
Stevens Institute of Technology
Hoboken, NJ, USA 07307
cshi5@stevens.edu

Jian Liu
Stevens Institute of Technology
Hoboken, NJ, USA 07307
jliu28@stevens.edu

Hongbo Liu
Indiana University Purdue University Indianapolis
Indianapolis, IN, USA 46202
hl45@iupui.edu

Yingying Chen
Stevens Institute of Technology
Hoboken, NJ, USA 07307
yingying.chen@stevens.edu

ABSTRACT

User authentication is a critical process in both corporate and home environments due to the ever-growing security and privacy concerns. With the advancement of smart cities and home environments, the concept of user authentication is evolved with a broader implication by not only preventing unauthorized users from accessing confidential information but also providing the opportunities for customized services corresponding to a specific user. Traditional approaches of user authentication either require specialized device installation or inconvenient wearable sensor attachment. This paper supports the extended concept of user authentication with a device-free approach by leveraging the prevalent WiFi signals made available by IoT devices, such as smart refrigerator, smart TV and thermostat, etc. The proposed system utilizes the WiFi signals to capture unique human physiological and behavioral characteristics inherited from their daily activities, including both walking and stationary ones. Particularly, we extract representative features from channel state information (CSI) measurements of WiFi signals, and develop a deep learning based user authentication scheme to accurately identify each individual user. Extensive experiments in two typical indoor environments, a university office and an apartment, are conducted to demonstrate the effectiveness of the proposed authentication system. In particular, our system can achieve over 94% and 91% authentication accuracy with 11 subjects through walking and stationary activities, respectively.

CCS CONCEPTS

•Security and privacy → Authentication;

ACM Reference format:
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 Mobihoc '17, Chennai, India, July, 2017,* 10 pages.
DOI: 10.1145/3084041.3084061

Mobihoc '17, Chennai, India

DOI: 10.1145/3084041.3084061

1 INTRODUCTION

In recent years, user authentication, the process of verifying the identity of a person who connects to private resources (e.g., accessing proprietary information and operating risky home appliances), has become increasingly vital due to the growing concern of user security and privacy leakage. For instance, unauthorized users may access the confidential documents or offices that only allow designated personnel or operate on private devices (e.g., computers) that always contain sensitive information. Furthermore, the emerging applications in smart homes/offices are also exploring the ability to distinguish different people and launch customized services accordingly, such as prohibiting children and elderly people to operate risky appliances (e.g., oven and dryer), adjusting room temperature/lighting conditions and recommending TV content. Such advancement of smart environments makes the user authentication process evolve to a broader context than traditional applications.

Traditional user authentication approaches mainly rely on either password [13], handwriting [3] or physiological biometrics such as fingerprints and iris [10, 12] to authenticate users. However, they usually require extra and dedicated devices installed before deployment. Other research studies reveal the behavioral features of users, such as key-press durations [17] during typing and mouse dynamics [26], could be applied to perform continuous user authentication. However, these approaches only work when the user operates the keyboard or mouse. Additionally, gait patterns [16] derived through mobile devices require users to carry additional devices when user authentication is performed. To support the evolving concept of user authentication, in this paper we explore a device-free approach that could perform user authentication through daily activities without requiring a user to carry any device. The basic idea is to exploit unique physical properties embedded in people's daily activities (e.g., entering an office with proprietary information, opening a refrigerator or cooking on a stove) to capture each person's physiological and behavioral characteristics to facilitate user authentication.

With the advent of Internet of Things (IoT) over the past decade, almost every electronics in indoor environments (such as smart refrigerator, smart TV, smart thermostat, home security system and wearable devices) are interconnected wirelessly. Because of this, the wireless connection among IoT devices provides rich web of reflected rays that spread every indoor corner. Although the wireless signal generated by IoT devices that are designed for many special

applications, it has the potential to capture human's unique physiological and behavioral characteristics inherited from people's daily activities when operating such devices (e.g., opening the refrigerator and entering the restricted office), which provides an appealing direction to differentiate each individual.

Recent studies [20, 24, 25] already show the success of using WiFi signals to capture gait patterns for user identification. However, these approaches only apply to a small group of people (i.e., 2 to 7) and are limited to walking people. They require either the users to walk through well-designed paths (e.g., clear Line of Sight (LoS) path between the WiFi devices) or have the WiFi transceivers placed close to each other, which is not practical in many real world scenarios. In contrast, we propose a device-free user authentication system relying on the existing WiFi signals generated by IoT devices to capture unique characteristics of each user inherited from both walking and stationary daily activities (e.g., working in front of the computer, operating a hot stove at home or entering an office with proprietary information).

In order to exploit human daily activities for user authentication, our device-free system should be able to recognize different types of daily activities and also differentiate each individual user if the same type of activity is performed. Thus, it is essential to derive representative wireless measurements to well capture the physiological (e.g., body shape, height, and weight) and behavioral characteristics (e.g., walking patterns, preferences when operating appliances) of each individual. Additionally, recognizing activities and identifying users require different granularity of abstractions from physiological and behavioral features. In general, activity recognition requires less feature granularity than human identification because coarse data representations are sufficient to recognize different types of activities with reasonable accuracy. Therefore, the designed system needs to have the capability to extract different levels of feature representations to perform activity recognition and further conduct human identification.

Toward this end, we propose to extract the representative features based on both amplitude and relative phase of Channel State Information (CSI) measurements in WiFi signals, which have the potential to reveal unique characteristics of different users. Furthermore, a three-layer deep neural network (DNN) model is developed to learn high-level abstractions of human physiological and behavioral characteristics for both activity recognition and human identification, which meets the hierarchical nature of our user authentication system involving different granularity levels of activity/human identification. In particular, the DNN scheme detects the activity type (i.e., stationary or walking) in the first layer and obtains the activity details (e.g., walking paths, opening a refrigerator) in the second layer. In the third layer, the model can learn the highest level non-linear abstractions from the representative features obtained from human activities and authenticate the user accordingly. Additionally, we also build one spoofing detection scheme based on support vector machine (SVM). Extensive experiments involving 11 subjects are conducted in both lab and apartment environments for testing accessing restricted areas and operating risky appliances. The results demonstrate that our device-free system can perform accurate user authentication through human daily activities, and is thus capable to facilitate many emerging applications (e.g., smart homes/offices and smart healthcare) in both corporation offices and residence areas. The main contributions of our work are summarized as follows:

- Our study shows that the existing WiFi signals generated by indoor IoT devices can be utilized to capture unique human physiological and behavioral characteristics and thereby authenticate users from their daily activities (i.e., both walking and stationary activities).
- Our proposed device-free system leverages a single pair of WiFi-enabled devices to extract both amplitude and relative phase from fine-grained channel state information (CSI) to facilitate accurate user authentication without the active participation of the users.
- We develop a deep learning based model to detect the uniqueness of human daily activities and capture the distinct WiFi fingerprints of different users. Our system is resilient to user spoofing attack by integrating with the SVM technique.
- Extensive experiments are conducted in both lab and apartment environments over a five-month period, and our system can achieve over 94% and 91% authentication accuracy through walking and stationary activities, respectively.

2 RELATED WORK

As a traditional way for user authentication, password based approaches [3, 13] require the users to remember either some secure texts or graphical patterns. Such authentication systems solely rely on the knowledge of password and thus are easily suffered from password stolen or shoulder surfing. Recent studies show success in exploring physiological biometrics such as fingerprints, iris, and facial information [4, 10, 12] to perform user authentication. These approaches, however, require dedicated equipments (e.g., fingerprint scanner or iris camera) before deployment.

To overcome the aforementioned weaknesses, the researchers seek for behavioral features for continuous user authentication. Some studies attempt to authenticate user by measuring users' behavioral characteristics such as key-press durations, multi-key latencies [17], angle preferences when operating a mouse [26]. However, these approaches require user active participation and can only work when the user operates the keyboard or mouse. Furthermore, Ren *et al.* leverage the acceleration readings from mobile devices to uniquely recognize user's walking gait patterns [16]. Ranja *et al.* propose to recognize the unique hallmarks [15] through wearable sensors readings when the users are operating home appliances. These authentication schemes require users to carry additional devices which may cause inconvenience for users.

Recently, WiFi based sensing attracts considerable attention from many researchers due to the prevalence of wireless signals in indoor environments. Previous studies propose to use WiFi signals to recognize human activities [22], estimate walking direction [23] and even monitor people's breathing rate while sleeping [11]. Furthermore, researchers demonstrate the possibility of utilizing wireless signals to perform user authentication. Existing studies [20, 24, 25] explore to capture human walking gait pattern and identify users in a small group by examining the CSI measurements. Specifically, Zhang *et al.* [25] extract a set of 10 features from CSI variations caused by human walking and uniquely identify each individual. Zeng *et al.* [24] propose a scheme which leverages WiFi characteristics to identify a person's steps and walking

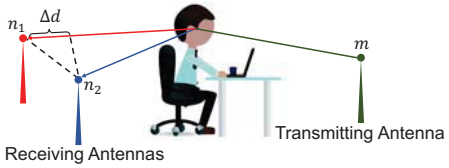


Figure 1: Relative phase produced by one transmitting antenna and two receiving antennas.

gait and further identify each user through the extracted CSI amplitude features. Additionally, Wang *et al.* [20] correlate movement speed of different body parts with WiFi spectrogram and perform gait pattern based user authentication. These approaches are limited to walking people and they require either the users to walk through well-designed paths (e.g., clear Line of Sight (LoS) path between the WiFi devices) or have the WiFi transceivers placed close to each other, which are not impractical in many scenarios. Moreover, WFID [7] performs device-free user authentication via characterizing the uniqueness of subcarrier-amplitude frequency (SAF) on CSI measurements when the users are standing, marching, and walking. Although activities of standing and marching are identified, the proposed SAF feature set does not capture the physiological and behavioral characteristics of users. Different from previous work, our system examines the WiFi signals and extracts unique physiological and behavioral characteristics inherited from people's daily activities including both walking activities (e.g., waking between rooms) and stationary activities (e.g., operating appliances) to differentiate each individual person. We exploit the unique individual characteristics from both amplitude and relative phase of CSI during people's daily activities. A deep learning based model is developed to learn deep representations and perform both activity recognition and user authentication, which is capable to facilitate many applications in both corporation offices and residential areas.

3 SYSTEM DESIGN

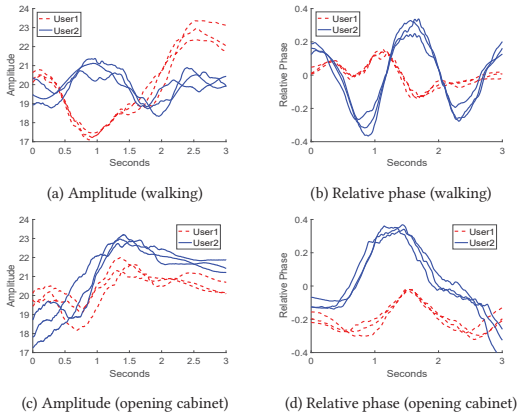
In this section, we first discuss some preliminaries of CSI and system design challenges, then we present the system overview.

3.1 Preliminaries

The prevalence of WiFi traffics in IoT environments, involving a multitude of smart devices and appliances (laptop, smart refrigerator, smart microwave oven and smart printer), can be exploited to capture the environmental changes induced by people's daily activities. Even for the same activity, different users exhibit subtle differences on the impact of wireless channel due to their unique physiological (e.g., body shape, height) and behavior characteristics (e.g., body moving). We are thus motivated to utilize the WiFi CSI measurements in IoT environment to monitor human activity and perform device-free user authentication.

Specifically, the fine-grained CSI describes how an OFDM signal propagates over multiple subcarriers between a pair of transmitter and receiver. It presents the combined effect of scattering, fading, and multi-path, which result in the distortion on the amplitude, phase and angle of arrival of the signal. Without loss of generality, the CSI between a pair of transmitting antenna m and receiving antenna n at the i_{th} subcarrier is defined as:

$$H_i^{m\leftrightarrow n} = |H_i^{m\leftrightarrow n}|e^{j\angle H_i^{m\leftrightarrow n}}, \quad (1)$$



(a) Amplitude (walking) (b) Relative phase (walking)

(c) Amplitude (opening cabinet) (d) Relative phase (opening cabinet)

Figure 2: CSI amplitude and relative phase of two users when walking or opening a cabinet.

where $|H_i^{m\leftrightarrow n}|$ and $\angle H_i^{m\leftrightarrow n}$ denote the amplitude and phase response, respectively. Previous studies have shown their success in utilizing CSI amplitude to perform activity recognition and identify users based on large scale body movements such as walking [20–22, 24]. However, the problem of authenticating user based on stationary activities (e.g., opening a cabinet) remains open due to the subtle movements of users.

To cope with the above challenge, we propose to utilize relative phase to capture the subtle changes of human physiological and behavioral characteristics. As the example depicted in Figure 1, the difference on signal path lengths, Δd, between two antennas (i.e., n_1 and n_2) varies as body moving and thereby results in relative phase shift (e.g., $m \leftrightarrow n_1$ and $m \leftrightarrow n_2$). The relative channel response at the i_{th} subcarrier can be formulated as:

$$\hat{H_i} = H_i^{m\leftrightarrow n_1}(H_i^{m\leftrightarrow n_2})^* = |\hat{H_i}|e^{j\angle\hat{H_i}}, \quad (2)$$

where $*$ denotes the complex conjugate, $\angle\hat{H_i} = -\frac{2\pi}{\lambda}\Delta d$ [9] is the relative phase value, Δd is the length difference of two transmitting paths, and λ is the signal wavelength. Given that cm-scale λ, relative phase is capable of capturing subtle movements for different users. Relative phase can also eliminate the impact of unpredictable offset on the absolute phase that is always hidden in the hardware control mechanism. Relative phase can avoid the significant unsynchronization of raw phase since the ever changing phase offset of transceivers can be eliminated in relative channel.

Figure 2 shows the extracted CSI amplitude and relative phase of a subcarrier over a 802.11n WiFi link over time when two users are walking along the same trajectory (3 rounds each) and opening a cabinet (3 rounds each), respectively in an office. We observe that both CSI amplitude and relative phase exhibit different variation trends between these two users, which confirms CSI is able to capture the unique physiological and behavioral characteristics of users. Additionally, for stationary activities (e.g., opening a cabinet), the difference on relative phase is more significant than that on amplitude, so it indicates the high sensitivity of the relative phase on capturing small-scale body movements.

3.2 Challenges

In order to authenticate people through people's daily activities via WiFi CSI measurements, a number of challenges need to be addressed.

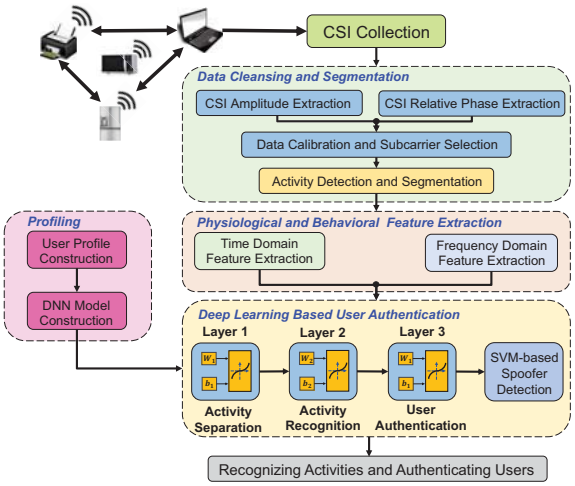


Figure 3: System Overview.

Uniqueness of Individual Characteristics. CSI measurements embedded in WiFi signals can be affected by user's subtle body movements while performing activities in the environments. Additionally, human's physical body (e.g., shape and height) also plays an important role to the multi-path effect, and signal interference. To authenticate users, the system needs to extract both unique physiological and behavior characteristics for each individual from the wireless signals.

System Robustness & Generality. The collected CSI measurements from real-world environments are usually noisy due to the continuous environmental changes and radio interference, etc. The system should be robust to capture distinguishable characteristics between users from such noisy channel measurements. Moreover, it is essential to design a general algorithm to extract representative features that keep and maintain user's physiological and behavioral characteristics through various daily activities.

Recognizing Activity & Identity Simultaneously. Recognizing activity and user identity simultaneously is very important in many smart home/office enabled applications. For instance, the system can prohibit a specific user (e.g., child) to watch TV at a specific time period. However, recognizing activities and identifying users require different granularity of features extracted from their activities.

3.3 System Overview

The basic idea of our system is to capture the unique physiological and behavioral characteristics inherited from human daily activities for user authentication leveraging WiFI signals. The users have habitual patterns on their behaviors, so the daily activities usually present high consistency for each individual [18]. As illustrated in Figure 3, our system takes as input CSI measurements from WiFi links between WiFi-enabled IoT devices (e.g., smart appliances), and then extracts both CSI amplitude and relative phase for each OFDM subcarrier for signal pre-processing. Unlike previous studies [20, 24, 25] which only utilize CSI amplitude, we explore relative phase to capture representative characteristics through user's daily activities. Given the amplitude and relative phase information, a band-pass filter is first deployed to eliminate the environmental interferences (e.g., reflected signals from furniture and walls) and ambient noises. We also propose a subcarrier selection algorithm to pick out the subcarriers with stable CSI measurements, which could represent reliable activity characteristics. Before performing features extraction, we examine the moving variance and related short time energy (STE) of the pre-processed data to determine the CSI segments, which capture the location changes for walking activities and body movement for stationary activities, respectively.

Next we will present the core components of our system, *Physiological and Behavioral Feature Extraction* and *Deep Learning Based User Authentication*. We perform activity recognition and user authentication based on the physiological and behavioral features extracted from CSI measurements, which characterize both human activity and identity uniqueness. The system extracts 6 time domain and 3 frequency domain features to capture both the physiological and behavioral characteristics of users such as height, shape and behavioral preference. Specifically, the time domain features, including maximum, minimum, mean, skewness, kurtosis and standard deviation, aiming to represent the extent of human movements and contour of human body, while the frequency domain features, including spectrogram energy, percentile frequency component, and spectrogram energy difference, are used to depict the fine-grained behavioral characteristics such as moving speed of torso and leg. All the above CSI-based features together provide a comprehensive and detailed representation for both walking and stationary activities. As far as we know, this is the first WiFi based user authentication scheme which authenticates users through both stationary activities and walking activities.

Finally, our system performs activity recognition and human authentication by building a three-layer deep neural network (DNN) model based on AutoEncoder [5]. Unlike previous authentication schemes based on high dimension feature sets and linear classification models (e.g., SVM), our DNN model learns non-linear physical and biometric abstractions which are computation efficient and are robust to small-scale input variations (e.g., the variations of features caused by the wearing changes of users). Particularly, we obtain the biometric abstractions with respect to single activity and authenticate the user based on the corresponding CSI activity segment. Figure 3 illustrates the functionality of each layer in our deep learning architecture for people authentication. In particular, the first level coarsely distinguishes the activity types (i.e., walking or stationary activity); the second layer exploits deep representations of the first layer and obtains the activity details such as walking trajectories and detailed stationary activity types (e.g., turning on a light, brushing teeth); and the third level obtains even deeper representation of the features and finally completes user authentication process. Additionally, our system is resilient to *user spoofing*, who either does not exist in legitimate user profiles or tries to mimic a legitimate user's activity, by using a SVM-based model with the generated DNN abstractions.

4 ACTIVITY SEGMENTATION AND FEATURE EXTRACTION

In this section, we first present how to perform data segmentation on the CSI measurements that reflect people's daily activities, and then extract effective features that capture unique physiological and behavioral characteristics of people from WiFi signals.

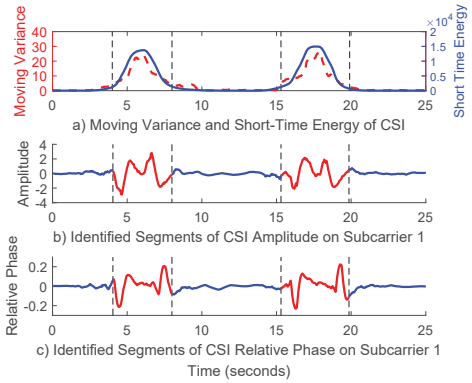


Figure 4: Illustration of activity detection and segmentation using CSI moving variance and short time energy.

4.1 Activity Detection and Segmentation

To ensure the reliability of the features extracted from the CSI measurements, the data calibration and subcarrier selection techniques are developed to mitigate the ambient noises and select the subcarriers with stable CSI measurements, respectively. The details are presented in Section 6.

As mentioned in preliminaries 3.1, both walking and stationary activities lead to the variations in wireless channel, resulting in changing CSI measurements. So we apply the short time energy (STE) upon CSI amplitude's moving variance to detect human activities, and then perform corresponding data segmentation. Moreover, stationary activities (e.g., opening a cabinet) usually involve relative smaller scale body movements than walking activities, which makes them even harder to be detected. We thus propose to examine STE which is more sensitive to subtle body movements because it is a summation of squared signals within a sliding window. We calculate STE within a sliding window as follows:

$$STE(t) = \sum_{n=1}^{N}(\sum_{k=1}^{K} v_k(t+n))^2, \tag{3}$$

where N is the length of the sliding window, $v_k(t)$ is the moving variance of CSI amplitude at the k_{th} subcarrier at time t.

Figure 4 (a) shows both CSI's moving variance and STE for two rounds of the same stationary activity (i.e., opening a cabinet). It is obvious that STE exhibits greater values when the activity occurs. Furthermore, we also found the peaks of the fluctuating part in STE always locate at the center of the activity duration. We are thus inspired to utilize a dynamic threshold, which is applicable to all types of activities, to perform activity detection and corresponding data segmentation. Specifically, the weight $w = 0.1$ according to our empirical study is deployed for the dynamic threshold calculation: $\tau = w*E$, where E is the maximum value of STE for different activities. We then search for the starting and ending points, t_s and t_e, of an activity by solving the following objective problem:

$$\begin{aligned} &\arg\min_{t_s,t_e} \quad t_s + t_e - 2t_m \\ &s.t., STE(t_s), STE(t_e) < \tau, STE(t_m) > \tau, \\ &\qquad t_s < t_m < t_e \end{aligned} \tag{4}$$

where t_m is an arbitrary time index in the middle of activity duration. Figure 4 (b) (c) show the segmented time series of CSI amplitude and relative phase of the 1_{st} subcarrier during two rounds of one stationary activity. The results demonstrate the efficiency of our activity detection and segmentation algorithm.

4.2 Physiological and Behavioral Feature Extraction

To capture the unique physiological and behavioral characteristics inherited from users' daily activities, it is essential to extract effective and reliable features from the CSI measurements. In particular, both time and frequency domain features based on CSI amplitude and relative phase information are examined to discriminate different users.

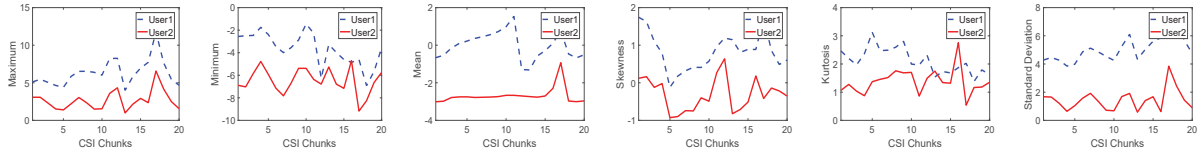
Time Domain Feature Extraction. In our system, 6 time domain features with respect to CSI amplitude and relative phase, including *maximum, minimum, mean, skewness, kurtosis* and *standard deviation*, will be extracted to characterize both human activity and identity uniqueness. In order to provide finer feature granularity for each individual activity, we first partition the CSI segment, where the activity characteristics are embedded, into l chunks of equal length, and each of them will extract all 6 time domain features as introduced above. We empirically set l as 20 and extract the 6 feature points in each chunk. Thus, 120 feature points could be extracted at each subcarrier. Figure 5 and 6 present the extracted time domain features for the same stationary activity (i.e., opening a cabinet) performed by two users based on amplitude and relative phase, respectively. We can find that these features are significantly different between two users. It encourages us to leverage these time domain features to capture human unique characteristics inherited from their daily activities.

Frequency Domain Feature Extraction. As indicated in previous work [21], CSI measurements in the frequency domain are able to reveal the speeds of WiFi path length changes caused by human movements. Therefore, besides the time domain features, we also extract the representative features in frequency domain to capture the users' behavioral characteristics.

To extract the features in the frequency domain, given a CSI segment, we first adopt short-time Fourier transform (STFT) to obtain the two-dimensional spectrogram for the CSI amplitude or relative phase of each subcarrier. More specifically, we calculate 1000 points FFT within a 100ms sliding window, shifting 50ms each time. We then use bicubic interpolation [8] to resize the spectrogram results into a matrix $M_{(i,j)}$ (i.e., 10-by-10 matrix) of fixed size, which maintains spectrogram in a consistent feature space for different activities. Next three frequency domain features on the top of $M_{(i,j)}$ are extracted: 1) *Spectrogram magnitude*: each element in the matrix $M_{(i,j)}$; 2) *Percentile frequency components (PFC)*: $PFC(i,n) = \frac{\sum_{j=1}^{n} M_{(i,j)}}{\sum_{j=1}^{10} M_{(i,j)}}$, where $n = 1 \ldots 10$, subjected to $PFC(i,n) \geqq 0.5$ and $PFC(i,n) \geqq 0.95$, indicate the moving speed of torso and leg [20]. 3) *Spectrogram difference between time windows:* the element-wise differences between two consecutive rows in $M_{(i,j)}$, which capture the acceleration or deceleration process of body movement. In total, we extract 210 frequency domain feature values from the CSI segment with respect to one specific activity.

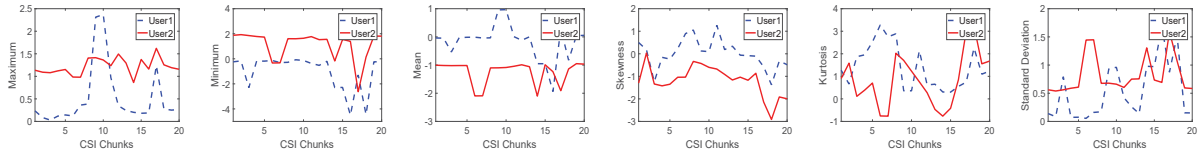
5 DEEP LEARNING BASED HUMAN AUTHENTICATION

In this section, we present the proposed deep learning based approach for both activity recognition and user authentication.



(a) Max (b) Min (c) Mean (d) Skewness (e) Kurtosis (f) STD

Figure 5: Time domain features of CSI amplitude over 20 chunks of one activity at the 1*st* subcarrier.



(a) Max (b) Min (c) Mean (d) Skewness (e) Kurtosis (f) STD

Figure 6: Time domain features of CSI relative phase over 20 chunks of one activity at the 1*st* subcarrier.

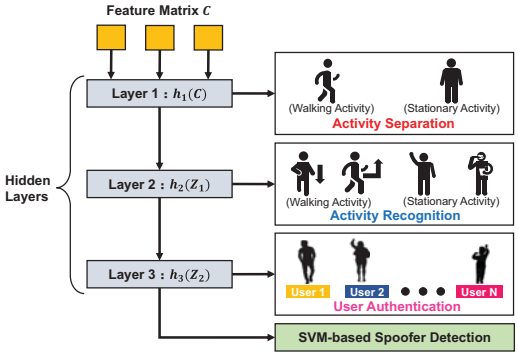


Figure 7: Deep learning architecture for user authentication.

5.1 Autoencoders Architecture

To perform activity recognition and further user authentication leveraging the extracted CSI features in subsection 4.2, we propose to develop a deep neural network (DNN) to extract high level abstractions from the extracted CSI features. As illustrated in Figure 7, a three-layer stacked autoencoder [19] is proposed based on deep neural network model. Given a set of CSI features C, we define h_i (i=1, 2, 3) as the activation functions to encode the input, which can be either CSI features or modeled abstractions (i.e., Z_1 or Z_2) of each layer, into a set of compressed representations, which is then fed into classification functions (e.g., SVM [2] or softmax function [1]) in each layer. Specifically, the proposed DNN network coarsely recognizes the activity type (i.e., stationary or walking) in the first layer and obtains the activity details (i.e., specific type of activities) in the second layer. We denote Z_1 and Z_2 as the outputs (i.e., high-level, complex abstractions as data representations) from the first two layers, respectively. The third layer identifies each individual user with a softmax function. Additionally, a SVM model with DNN abstractions is integrated to ensure spoofing attack resilient in our system.

5.2 Per Layer Abstraction Extraction

Each layer of the DNN consists of an autoencoder network of neural units which learns a set of compressed representations from input features through unsupervised pre-training. Such compressed representations are able to characterize the physiological and behavioral uniqueness for different users. Previous work has been proved that such pre-training process is significantly helpful in classification tasks [5]. In addition, the autoencoder based DNN networks can learn abstractions with different granularity which can be leveraged for activity recognition and human identification. Particularly, the non-linear neural units in the hidden layer of autoencoder map the input X into a set of abstractions, Z as follows:

$$Z = \sigma(wX + b), \tag{5}$$

where $\sigma()$ is a logistic sigmoid function defined as $\sigma(z) = \frac{1}{1+e^{-z}}$, and w and b represent the weight and bias of the autoencoder, respectively. The autoencoders are trained in an unsupervised manner with the objective to minimizes the error when recovering the input X from Z. Specifically, the error cost function is defined as follows:

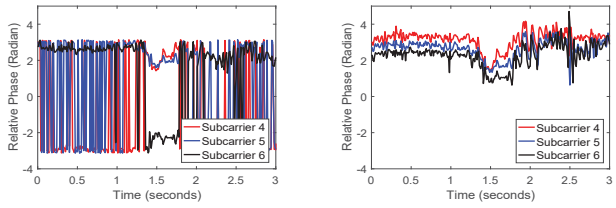
$$ERR(X, X') = \frac{1}{N}\sum_{n=1}^{N}(X_n - X_n{}')^2 + \lambda \times \Omega_{weights} + \beta \times \Omega_{sparsity}, \tag{6}$$

where N is the number of training samples, X' is the recovered X using Z from a decoder function, and $\Omega_{weights}$ and $\Omega_{sparsity}$ are the parameters of L_2 regulariser and spare regulariser, respectively, which prevent low neuron outputs of the autoencoder. We use λ and β to denote the coefficient of L_2 regulariser and sparse regulariser.

5.3 Activity Recognition and User Authentication

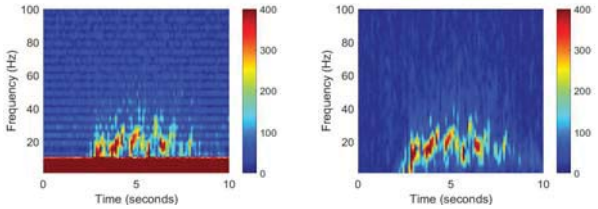
Given the three layers of autoencoders, a hierarchical abstraction learning architecture is constructed by stacking one layer of autoencoder on the top of another. Previous work [19] found that higher level feature abstractions are more stable and robust to small-scale input variations, which meets the hierarchical requirements of our system. Additionally, the three layer DNN itself can only derive compressed representations of physiological and behavioral characteristics, so we still need a softmax function [1] in each layer to complete the activity recognition and user authentication process in a hierarchical order. Specifically, we define the softmax function as follows:

$$P(L_k|Z) = \frac{P(Z|L_k)P(L_k)}{\sum_{j=1}^{K} P(Z|L_j)P(L_j)}, \tag{7}$$



(a) Raw CSI Relative Phase (b) Unwrapped Relative Phase

Figure 8: Correcting relative phase to eliminate phase offset using unwrapping function.



(a) Before data calibration (b) After data calibration

Figure 9: CSI amplitude spectrogram of at the 1_{st} subcarrier before and after data calibration.

where $P(L_k|Z)$ denotes the posterior probability of class label L_k given an abstraction Z, and $P(L_k)$ represents the prior of the same class. We use $P(Z|L_k)$ to denote likelihood of the abstraction Z given label L_k. In addition, the equation is constrained by $0 < P(L_k|Z) \leq 1$ and $\sum_{k=1}^{K} P(L_k|Z) = 1$. The outputs of each softmax function characterize the probability distribution over K profiled classes (e.g., activity type or user identity), and the abstraction Z will be classified as class k, which satisfies $k = argmax_{k \in K} P(L_k|Z)$.

5.4 SVM based Spoofer Detection

Besides three-layer DNN, we also adopt a SVM model [2] to determine whether the activity the user performed matches one of the legitimate user profiles. Particularly, we propose to utilize one-class support vector machine with Gaussian kernel for detecting the user spoofing, who either does not exist in legitimate user profiles or tries to mimic a legitimate user's activity. We first construct an one-class SVM model for each of the legitimate users based on the high level abstractions from DNN network. We then derive a class score S_u, which compares the similarity between the DNN abstractions of each testing sample and the support vectors of the profile of user u:

$$S_u(Z) = \sum_{i}^{N_u} k(Z_{u,i}, Z) + b_u, \tag{8}$$

where Z is a sample abstraction, $Z_{u,i}$ is the ith support vector of the user u, $k()$ represents the Gaussian kernel function, and b_u is the function bias. Greater value of the class score S_u represents that the testing sample has the less distance to the the support vectors of user u. An empirically set threshold η thus is used to detect possible spoofers. The testing sample would be determined as spoofer/attacker if the derived class scores (i.e., S_u) are less than η from all the legitimate user profiles.

6 DATA CALIBRATION & SUBCARRIER SELECTION

In this section, we introduce how to ensure the reliability of the extracted CSI amplitude and relative phase from the noisy wireless signal readings.

6.1 Data Calibration

To ensure reliable feature extraction, we preprocess the raw CSI measurements with phase unwrapping and a band-pass filtering techniques, which are effective on eliminating the environmental interferences (e.g., reflected signals from static objects) and ambient noises. Specifically, we first eliminate the relative phase error caused by the phase offset on each subcarrier. As shown in Figure 8 (a), the time series of raw relative phase at three subcarriers (i.e., subcarrier 4, 5 and 6) have obvious discontinuities between consecutive packets when the relative phase value is close to $\pm\pi$. To eliminate such discontinuity, a $\pm 2\pi$ is added to the relative phase of the later packet if the absolute phase difference of two consecutive packets is greater than or equal to π. Figure 8 (b) shows the corresponding relative phase streams after phase calibration.

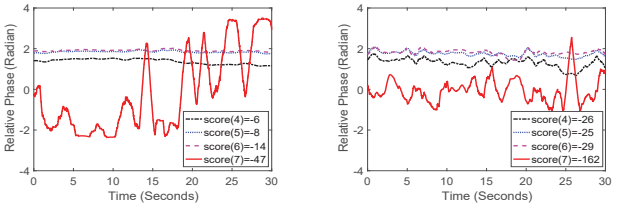
Besides the phase offset, the amplitude and relative phase in CSI measurements are also easily affected by the low frequency interference (i.e., reflected signals from static objects) and high frequency noise. In order to eliminate the above impacts while preserving the user physiological and behavioral characteristics in the CSI measurements, a bandpass butterworth filter [14] is adopted in the data calibration. Previous work [21] found that the frequency range of most human activities including running in a fast speed exhibit CSI frequency components less than $300Hz$. In our indoor home/office scenarios, we thus adopt a relative low frequency band-pass butterworth filter (i.e., with passing band $5Hz$-$100Hz$) to effectively remove both low and high frequency components from the spectrum. Given the example scenario where a person walks in the room between 2 and 7 sec, Figure 9 (a) shows the spectrogram of the corresponding time series of CSI amplitude at a subcarrier (i.e., subcarrier 1). We can observe that the spectrogram exhibits extremely high energy level in the low frequency band (i.e., $< 10Hz$) even the person remains static. As the spectrogram after band pass filtering shown in Figure 9 (b), we can observe that the CSI amplitude pattern caused by human walking is still preserved while irrelevant frequency components are removed.

6.2 Subcarrier Selection

Our preliminary study finds that the CSI measurements of several subcarriers are more sensitive to ambient noise. To ensure the reliability of CSI measurements for later processing, we propose a new subcarrier selection method to determine the noise resilient subcarriers from the CSI measurements. The CSI measurements at neighboring subcarriers are usually highly correlated, however such correlation could be destroyed by heavy noises on the subcarriers. To eliminate the negative effects caused by the unstable subcarriers, a covariance based scoring function is defined to assess each subcarrier's correlation level with its neighboring subcarriers as follows:

$$score(i) = \sum_{n=1}^{N} \sum_{j=i-\frac{k}{2}}^{i+\frac{k}{2}} \frac{cov_{i,j}(t) - |cov_{i,j}(t)|}{2}, \tag{9}$$

where N is the number of non-overlapped $1s$ length time windows being divided in the short phase, k is the number of its close-by subcarriers being compared, and $cov_{i,j}$ denotes the covariance value



(a) Empty room (b) Subject walking in the room

Figure 10: Detecting noisy subcarriers by using a covariance based scoring function: Subcarrier 7, which has the lowest score, is not a stable subcarrier that can be used in the system.

between the CSI relative phase at the i_{th} and j_{th} subcarriers. Figure 10 presents an example showing the scores of 4 subcarriers (i.e., subcarrier 4, 5, 6 and 7) based on the CSI measurements collected in a 30s short period. We can observe that the CSI measurements of subcarrier 7 keep fluctuating in both empty room and human walking cases, so it implies the instability of the subcarrier 7 is not caused by human movements. As a result, the subcarrier 7 has the lowest score, indicating it has the lowest correlation with its adjacent subcarriers. In our system, to remove the noisy subcarriers while keeping consistent dimension in the feature space, we choose the top 20 subcarriers with the highest scores to collect CSI measurements.

7 PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed system on activity recognition and user authentication in both a university office and an apartment.

7.1 Experimental Methodology

Devices and Network. We emulate the WiFi network in IoT environments with two commercial laptops equipped with 802.11n WiFi NICs (i.e., Intel 5300 NICs). Specifically, we deploy a Dell $E6430$ laptop as transmitter and a Lenovo $T61$ laptop as receiver. Both of the transmitter and receiver run Ubuntu 14.04 operating system with the 4.2.0 kernel for measuring CSI over 30 subcarrier groups [6]. We extract the CSI amplitude on the link between the main antenna pair (i.e., 1_{st} antenna in both transmitter and receiver), and compute the relative phase of CSI between the two links from the transmitter's main antenna to the first two antennas on the receiver. In addition, the packet transmission rate is fixed at 1000 *pkts/s* to enable high resolution analysis in the frequency domain.

Environments and Activities. The proposed system is evaluated in both a university office and an apartment with the size of $26ft \times 14ft$ and $36ft \times 22ft$, respectively. Figure 11 shows the experimental setups involving two laptops to emulate as IoT-enabled devices (e.g., smart refrigerator, smart TV and smart thermostat) and generate WiFi traffics.

A total of 8 walking activities and 8 stationary activities (30 rounds for each) are performed by 11 and 5 volunteers in these two indoor environments, respectively. Due to the functionality differences of the two environments, we choose different yet still typical stationary activities in the two environments. The details of the activities are listed in Table 1, and the locations of stationary activities and walking trajectories are also shown in Figure 11. In

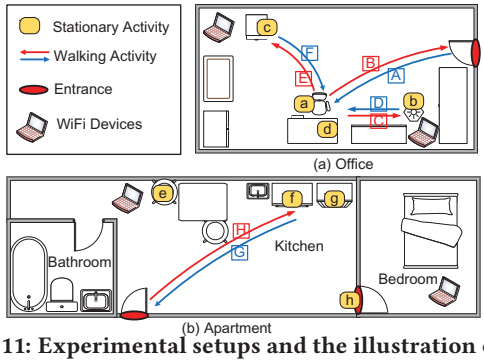


Figure 11: Experimental setups and the illustration of activities in an office and an apartment.

Table 1: Detailed daily activities performed.

Code	Walking activity	Code	Stationary activity
A	Entrance→Seat	a	Working (i.e., typing keyboard)
B	Seat→Entrance	b	Turning on the light
C	Seat→Light Switch	c	Opening the cabinet
D	Light Switch→Seat	d	Fetching documents
E	Seat→Cabinet	e	Eating at the table
F	Cabinet→Seat	f	Opening the microwave oven
G	Entrance→Kitchen	g	Opening the refrigerator
H	Kitchen→Entrance	h	Opening the door

total, we collect 3336 activity segments performed by 11 subjects in the office environment, and 834 activity segments performed by 5 subjects in the apartment. Unless mentioned otherwise, half of the collected data-set (i.e., 15 rounds of each activity per subject) is used for training the DNN model, and the rest of data is used for testing the system performance.

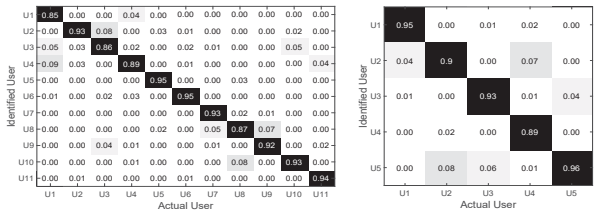
Classification Strategies. Our system first collects CSI measurements through WiFi scanning while people are performing daily activities. Then, we associate each activity with the corresponding segment of CSI measurements labeled as the activity profiles for the legitimate users. Given the activity profiles, we extract both time and frequency domain features from the CSI measurements for any unknown activities, and feed them to the DNN model for gesture recognition and user identification.

Evaluation Metrics. To evaluate our system performance, we define the following metrics:

- **Confusion Matrix.** Each column in the confusion matrix indicates the ground truth of an identity/activity and each row represents the classified identity/activity in our system. Each entry in the matrix represents the percentage of correctly classified identify/activity.
- **Identification Accuracy.** The percentage of the human identity/activity correctly recognized by our system.
- **Spoofer detection accuracy.** The percentage of the spoofing activities being correctly detected by our system.

7.2 User Authentication Performance

We first present the performance of the proposed system on user identification in both office and apartment environments. As shown in Figure 12 (a), we observe that, our system achieves over 92% user identification accuracy for 7 out of 11 users and the average accuracy is 91.2% with a standard deviation of 3.67% in the office environment. Figure 12 (b) gives the confusion matrix for the



(a) Office (b) Apartment

Figure 12: Performance of deep learning based user authentication.

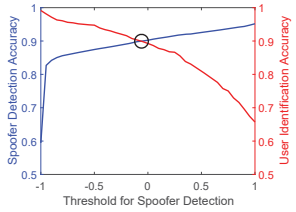


Figure 13: Performance of spoofer detection.

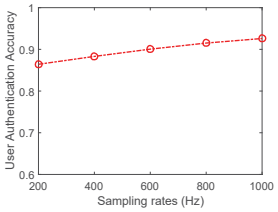


Figure 14: Impacts of sampling rate.

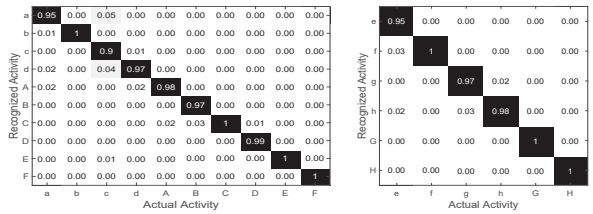
user identification in the apartment. Our system achieves over 90% identification accuracy for 4 of the users. The average user identification accuracy is 92.4% with a standard deviation of 3.49%. We have comparable high accuracies on user authentication in the two different environments, and thereby confirm the effectiveness and reliability of the proposed system on user identification.

7.3 Spoofer Detection Performance

We next evaluate our SVM based spoofer detection algorithm in the office environment, where 3 of the 11 users are legitimate users and the other 8 users act as spoofers. Both the spoofer detection and legitimate user identification accuracy are presented in Figure 13 with varying spoofer detection threshold η (i.e., $[-1, 1]$). As the figure indicates, the spoofer detection accuracy grows with the positive class score. This is because as the threshold increases, the region decision encircled by the decision boundary becomes smaller and more CSI samples from spoofer can be excluded. We also find that the legitimate user detection ratio decays with greater threshold. So we seek for a tradeoff to maintain both high spoofer detection accuracy and high legitimate identification accuracy. According to Figure 13, an appropriate threshold is determined by the intersection of spoofer detection ratio curve and legitimate user detection ratio curve. Therefore, we can achieve the accuracy as high as 89.7% for both spoofer and legitimate detection, which validates the robustness of the proposed system on user authentication under the spoofing attack.

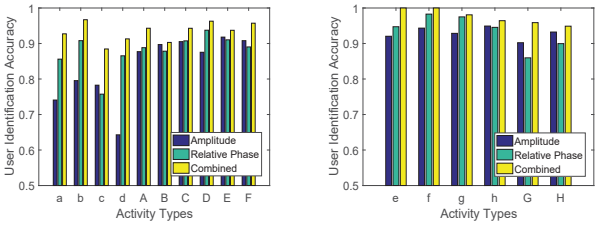
7.4 Activity Recognition Performance

We further examine the activity recognition performance in the DNN model. Figure 15 depicts the confusion matrix for activity recognition (i.e., outputs of DNN layer 2) in both office and apartment environments. The average activity recognition accuracies for both stationary and walking activities are as high as 97.6% and 98.3% in the office and apartment, respectively. Further it is encouraging to find that DNN model achieves similar accuracy for both stationary and walking activities. The slight difference on the recognition accuracy between the two types of activities is caused



(a) Office (b) Apartment

Figure 15: Performance of deep learning based activity recognition.



(a) Office (b) Apartment

Figure 16: Comparison of features used in deep learning based user authentication under different activities.

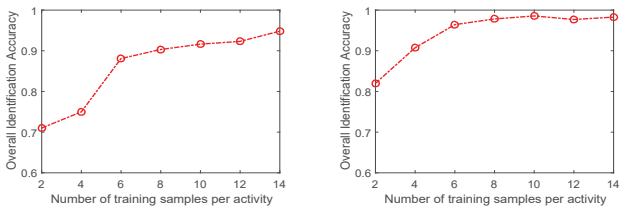
by the limited resolution of WiFi signals on capturing small scale body movements for stationary activities. Overall, the proposed DNN model is highly effective on recognizing different types of activities.

7.5 Impact of Various Factors

Impact of Sampling Rate. In order to validate that our user authentication scheme can work under various sampling frequencies of WiFi enabled IoT devices, we evaluate our system under different frame rates. We show the average user authentication accuracy of office and apartment under different sampling rates in Figure 14. We can observe that our system can maintain high accuracy across different frame rates from $200Hz$ to $1000Hz$. Particularly, the authentication accuracy is still over 86% even for the low sampling rates such as $200Hz$ and $400Hz$. The above observations confirm that our system can be applied on IoT devices with different sampling capabilities.

Feature Comparison. To further analyze the impact of different features on the system performance, we compare the authentication performance using different kinds of CSI features in both time and frequency domains: *Amplitude, Relative phase* and all of these features (i.e., *Combined*). We present the comparison results of user authentication accuracy in Figure 16 for both office and apartment environments. Figure 16 (a) shows that CSI relative phase features have relative higher user identification accuracy for stationary activities (e.g., a, b, c, d) comparing to the amplitude feature. This is primarily because relative phase exhibits higher sensitivity on capturing small-scale body movements. In addition, we also find in both Figure 16 (a) and (b) that the combined features of both CSI amplitude and relative phase achieve the best performance, indicating the combining features can provide finest features to distinguish individual subject.

Impact of Training Size. The DNN model needs to build CSI profiles for each activity or each individual before performing activity recognition and human authentication. It is necessary to study



(a) User identification (b) Activity recognition

Figure 17: User authentication and activity recognition performance under different training sizes.

the impact of training size on the system performance. Here we define the training size as the number of training samples for each activity or for each individual. As shown in both Figure 17 (a) and (b), our system can achieve consistently high accuracy on user identification and activity recognition with different training sizes. Especially, our deep learning model maintains over 90% accuracy on user identification and activity recognition even with the training size of 4. The above results show that our system has minimum requirement on building the CSI profile while ensure remarkable performance.

8 CONCLUSION

As the proliferation of Internet of things (IoT), the prevalence of wireless connections among IoT devices provides the opportunity to authenticate users through examining the wireless signal characteristics inherited from daily activities. In this paper, we propose a device-free user authentication system by extracting unique physiological and behavioral characteristics embedded in human daily activities captured by the fine-grained Channel State Information (CSI). Our system takes one step forward to support the extended user authentication concept in not only preventing unauthorized users to access restricted information but also identifying users for customized services (e.g., prohibiting a kid to operate a hot stove) in both corporate and home environments. We find that both amplitude and relative phase available in CSI readings are impacted by the environmental changes caused by human activities in different scales. To extract meaningful patterns from noisy CSI measurements, we design data calibration and subcarrier selection algorithms to filter out various noises while preserve human physiological and behavioral characteristics. A deep learning based user authentication mechanism is developed leveraging the extracted CSI features in both time and frequency domains to accurately identify each individual. This is a first authentication system that has the capability of actuating daily activities for human authentication without active user participation nor attaching any devices to users. We show that the proposed system is resilient to spoofing attacks when integrating the feature abstractions from the deep learning model with the SVM classifier.

9 ACKNOWLEDGEMENT

This work was supported by the National Science Foundation Grant CNS1514436.

REFERENCES

[1] Christopher M Bishop. 2006. Pattern recognition. *Machine Learning* 128 (2006).
[2] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.
[3] Rachna Dhamija and Adrian Perrig. 2000. Deja Vu-A User Study: Using Images for Authentication. In *in Proceedings of the 9th conference on USENIX Security Symposium (SSYM)*, Vol. 9. 4–4.
[4] Benoit Duc, Stefan Fischer, and Josef Bigün. 1999. Face authentication with Gabor information on deformable graphs. *IEEE Transactions on Image Processing (IEEE TIP)* 8, 4 (1999), 504–516.
[5] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio. 2010. Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research* 11 (2010), 625–660.
[6] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* (2011), 53–53.
[7] Feng Hong, Xiang Wang, Yanni Yang, Yuan Zong, Yuliang Zhang, and Zhongwen Guo. 2016. WFID: Passive Device-free Human Identification Using WiFi Signal. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (ACM MobiQuitous)*. 47–56.
[8] Robert Keys. 1981. Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing* 29, 6 (1981), 1153–1160.
[9] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using WiFi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (ACM SIGCOMM)*. 269–282.
[10] Ajay Kumar and Arun Passi. 2010. Comparison and combination of iris matchers for reliable personal authentication. *Pattern recognition* 43, 3 (2010), 1016–1026.
[11] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking Vital Signs During Sleep Leveraging Off-the-shelf WiFi. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM MobiHoc)*. 267–276.
[12] Cástor Mariño, Manuel G Penedo, Marta Penas, María J Carreira, and F Gonzalez. 2006. Personal authentication using digital retinal images. *Pattern Analysis and Applications* 9, 1 (2006), 21–33.
[13] Robert Morris and Ken Thompson. 1979. Password security: A case history. *Commun. ACM* 22, 11 (1979), 594–597.
[14] Lawrence R Rabiner, Bernard Gold, and CK Yuen. 1978. Theory and application of digital signal processing. *IEEE Transactions on Systems, Man, and Cybernetics* 8, 2 (1978), 146–146.
[15] Juhi Ranjan and Kamin Whitehouse. 2015. Object hallmarks: Identifying object users using wearable wrist sensors. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*. 51–61.
[16] Yanzhi Ren, Yingying Chen, Mooi Choo Chuah, and Jie Yang. 2013. Smartphone Based User Verification Leveraging Gait Recognition For Mobile Healthcare Systems. In *Proceedings of the Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*. 149–157.
[17] Kenneth Revett. 2009. A bioinformatics based approach to user authentication via keystroke dynamics. *International Journal of Control, Automation and Systems* 7, 1 (2009), 7–15.
[18] Burrhus Frederic Skinner. 1953. *Science and human behavior*. Simon and Schuster.
[19] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research* 11 (2010), 3371–3408.
[20] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*. 363–373.
[21] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (ACM MobiCom)*. 65–76.
[22] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking (ACM MobiCom)*. 617–628.
[23] Dan Wu, Daqing Zhang, Chenren Xu, Yasha Wang, and Hao Wang. 2016. WiDir: walking direction estimation using wireless signals. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*. 351–362.
[24] Yunze Zeng, Parth H Pathak, and Prasant Mohapatra. 2016. WiWho: WiFi-based Person Identification in Smart Spaces. In *Proceedings of 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IEEE IPSN)*. 1–12.
[25] J Zhang, B Wei, W Hu, and S Kenhere. 2016. WiFi-ID: Human identification using WiFi signal. In *Proceedings of International Conference on Distributed Computing in Sensor Systems (IEEE DCOSS)*. 75–82.
[26] Nan Zheng, Aaron Paloski, and Haining Wang. 2011. An efficient user verification system via mouse movements. In *Proceedings of the 18th ACM conference on Computer and Communications Security (ACM CCS)*. 139–150.