Secure User Authentication Leveraging Keystroke Dynamics via Wi-Fi Sensing

Yu Gu, *Senior Member, IEEE*, Yantong Wang, Meng Wang, Zulie Pan, Zhihao Hu, Zhi Liu, *Senior Member, IEEE*, Fan Shi, and Mianxiong Dong, *Senior Member, IEEE*

***Abstract*—User authentication plays a critical role in access control of a man-machine system, where the knowledge factor, such as a personal identification number, constitutes the most widely used authentication element. However, knowledge factors are usually vulnerable to the spoofing attack. Recently, the inheritance factor, such as fingerprints, emerges as an efficient alternative resilient to malicious users, but it normally requires special equipment. To this end, in this article, we propose WiPass, a device-free authentication system only leveraging the pervasive Wi-Fi infrastructure to explore keystroke dynamics (manner and rhythm of keystrokes) captured by the channel state information to recognize legitimate users while rejecting spoofers. However, it remains an open challenge to characterize the behavioral features hidden in the human subtle motions, such as keystrokes. Therefore, we build a signal enhancement model using Ricean distribution to amplify user keystroke dynamics and a hybrid learning model for user authentication, which consists of two parts, i.e., convolutional neural network based feature extraction and support vector machine based classification. The former relies on visualizing the channel responses into time-series images to learn the behavioral features of keystrokes in energy and spectrum domains, whereas the latter exploits such behavioral features for user authentication. We prototype WiPass on the low-cost off-the-shelf Wi-Fi devices and verify its performance. Empirical results show that WiPass achieves on average 92.1% authentication accuracy, 5.9% false accept rate, and 6.3% false reject rate in three real environments.**

***Index Terms*—Behavioral features, channel state information (CSI), convolutional neural network (CNN), Ricean distribution, user authentication.**

I. Introduction

With the digital and network technology integrating into our daily lives, user authentication constitutes the first and most fundamental step in current man-machine systems to protect our increasing private and sensitive data. For our current computer systems, the most commonly used type of authentication element is the knowledge factor, i.e., something the user knows, such as a personal identification number (PIN) or a password.

However, the blossom of information technology dramatically increases the burden of maintaining such knowledge. For example, a survey conducted at 2006 involving one million respondents indicated that each user had over 25 different accounts on average, and that number rapidly grew to 80 at 2018 [1]. As a result, over one-fifth users adopt a unified password for all account, which sharply increases the system vulnerability to malicious attacks and may result in serious consequences. It is reported that 543 million employee credentials for Fortune 1000 companies are circulating on commonly used underground hacking forums, a 29% increase from 2020 [2].

Consequently, the inheritance factor, i.e., something the user is or does, emerges recently with the mature of biometrics recognition (face or fingerprint) as an promising alternative, especially for the widespread mobile devices. However, for the conventional desktop computers, we need to extend them with extra devices, such as fingerprint reader or facial cameras, raising the cost and availability issues and, thus, possibly offsetting its merits in real world. Moreover, unlike mobile devices, which are personal in essence, the privacy issue of computers in public cannot be overlooked.

Naturally, one question arises: *Is it possible to explore the convenient, safer, and unforgotten inheritance factors for user authentication but without any additional device?* To this end,

Manuscript received June 25, 2021; revised August 11, 2021; accepted August 20, 2021. Date of publication September 1, 2021; date of current version December 27, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61772169, in part by the Basic Business Expenses of Central Public Welfare Scientific Research Institutes of the Chinese Academy of Medical Sciences under Grant 2020-JKCS-002, in part by the Key Research and Development Plan of Anhui Province under Grant 202004b11020018, and in part by JSPS KAKENHI under Grant JP20F20080. Paper no. TII-21-2632. *(Corresponding author: Yantong Wang.)*

Yu Gu, Yantong Wang, and Meng Wang are with the Anhui Province Key Laboratory of Affective Computing and Advanced Intelligence Machine and the School of Computer and Information, Hefei University of Technology, Hefei 230009, China (e-mail: hfut_bruce@hfut.edu.cn; wangyantong912@mail.hfut.edu.cn; mengw512@mail.hfut.edu.cn).

Zulie Pan, Zhihao Hu, and Fan Shi are with the College of Electronic Engineering, National University of Defense Technology, Hefei 230047, China, and also with the Anhui Province Key Laboratory of Cyberspace Security Situation Awareness and Evaluation, Hefei 230037, China (e-mail: panzulie17@nudt.edu.cn; huzhihao@nudt.edu.cn; shifan17@nudt.edu.cn).

Zhi Liu is with the Department of Computer and Network Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: liu@ieee.org).

Mianxiong Dong is with the Department of Sciences and Informatics, Muroran Institute of Technology, Muroran 0508585, Japan (e-mail: mx.dong@csse.muroran-it.ac.jp).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2021.3108850.

Digital Object Identifier 10.1109/TII.2021.3108850

we propose WiPass, a password-free and device-free user authentication system leveraging user behaviormetric with the pervasive commodity Wi-Fi infrastructure. More specifically, the key idea of WiPass is to explore the keystroke dynamics, the manner, and rhythm in which an individual types characters on a keyboard, to identify users. WiPass is password-free since we rely on the behaviormetric of keystrokes rather than the password itself for user authentication. Hence, we can use a unified and easy-to-remember phrase for all users rather than a complex password for each user. WiPass is device-free since we exploit the existing Wi-Fi infrastructures for sensing keystrokes instead of mounting any additional devices on the subject. Also, WiPass needs no participation or cooperation of the subject. Moreover, wireless signal is imperceptible for users, easing their privacy concerns.

Wi-Fi signal, either running at 2.4 or 5 GHz, will scatter and bound off once upon the human body, creating new propagation paths. Therefore, there exists a recent trend of exploiting such multipath effect on Wi-Fi channel state information (CSI) for device-free human activity recognition [3], [4]. On one hand, state-of-the-art research even enables subwavelength sensing for subtle and periodic activities, such as respiration and heart beating [5]. On the other hand, keystrokes lack such periodic features and, thus, remains an open problem for Wi-Fi sensing.

In particular, designing WiPass encounters following two critical challenges at the signal and data layer, respectively.

1) *Signal Processing*: How to capture the keystroke-induced traces in Wi-Fi signal? Keystrokes are centimeter-level activities, which only slightly affect the Wi-Fi radio wave. Such effect, if not processed properly, would be easily covered by ambient noise.
2) *Data Processing*: How to recover the behaviormetric from the traces for user authentication? Different persons have different manners and rhythms of keystrokes and, thus, an accurate and rich characterization of such behaviormetrics is required.

To address the aforementioned challenges, WiPass is designed with the following two tightly coupled modules.

Ricean fading based signal enhancement. This module exploits Ricean fading to enhance the keystroke-induced traces in Wi-Fi signal. More specifically, we build an analytical model based on the Ricean fading to mathematically describe the process of keystrokes affecting the Wi-Fi signal. This model proves that the keystroke-induced traces all hide in the non-line-of-sight (NLoS) components of Wi-Fi propagation paths. Therefore, by intentionally blocking the line-of-sight (LoS) components, we can raise the sensitivity of Wi-Fi signal to keystrokes. Hence, we are able to generate a filtered data profile in terms of Wi-Fi CSI containing enhanced descriptions about the keystroke-induced traces for the next module to process.

Deep learning based user authentication. This module leverages deep learning to extract behaviormetrics of keystrokes from the enhanced Wi-Fi CSI for user authentication. In particular, we first visualize Wi-Fi CSI into time-series images so that mature feature extraction techniques in computer vision (CV) can be used. Then, we design a convolutional neural network (CNN), well-reputed for its generalization ability, to extract behaviormetric features of keystrokes in both energy and spectrum domains. Finally, we employ a support vector machine (SVM) classifier, widely approved for handling small-scale features, to recognize legitimate users while rejecting spoofers.

We realize WiPass with low-cost off-the-shelf Wi-Fi devices and carry out extensive real-world experiments, where WiPass has achieved 92.1% authentication accuracy (AA), 5.9% false accept rate (FAR), and 6.3% false reject rate (FRR) on average with 15 volunteers in three different environments, i.e., a 8.85×8.85 m^2 conference room, a 13.62×44.25 m^2 lobby, and a laboratory with the same size of the lobby. Moreover, WiPass is shown to be robust to various key system parameters, such as the unified password, positions of the keyboard, and experimental sites. Experiments also verified that a simple and easy-to-remember password is as effective as a complex one, since WiPass relies on keystroke dynamics rather than password itself for authentication.

The rest of this article is organized as follows. Section II reviews some representative prior works for user authentication. Section III introduces Ricean fading based signal enhancement. The design details of WiPass are presented in Section IV, followed by experimental evaluation in real world in Section V. Finally, Section VI concludes this article.

II. Related Work

WiPass roots in prior work on keystroke dynamic based user authentication and Wi-Fi-based user authentication, as summarized in Table I. We will briefly introduce them in this section.

A. Keystroke Dynamics Based User Authentication

User authentication, the act of proving the identity of certain users, constitutes a fundamental problem in the man-machine systems. In general, there exist three authentication elements, i.e., possession factor (something the user has, such as an ID card), knowledge factor (something the user knows, such as a password), and inheritance factor (something the user is, such as fingerprints, or does, such as *keystrokes*).

Keystroke dynamics traces back to the early days of the telegraph, when operators develop distinctive manners and rhythms of typing that can be used to differentiate them. Such dynamics was named as "fist of the sender," which has been found quite effective in identifying the source of Morse code during World War II [20]. Since then, keystroke dynamics has been successfully leveraged for user authentication with the evolving of man-machine systems, e.g., from desktops with the 107-key standard keyboard [21] to ATMs [22] and mobiles [23] mounted with the ABC-9-key.

But why keystroke dynamics constitutes a qualified element for authentication remains unclear until Joyce and Gupta in 1990 revealed similar neurophysiological factors, which make written signatures unique also exist in keystroke dynamics for the first time [24]. Joyce and Gupta picked up keystroke latencies (i.e., time between successive keystrokes) for authentication [24], which enlightened follow-up works to various other hand-crafted features, such as keystroke duration (i.e., length of

TABLE I
LITERATURE REVIEW

Keystroke Dynamics based User Authentication via Different Signals			
	Signal	Algorithm	Performance
Roth 2014 [6]	Vision	Bag of Multi-dimensional Phrases	EER: 4.9%; TPR: 99.75% FPR: 0.001%
Ayotte 2019 [7]	Vision	Distance Metric Fusion	EER: 3.6%
Lv 2006 [8]	Pressure Sensors	DTW + Hybrid distances	EER: 1.41%
Giuffrida 2014 [9]	Pressure Sensors	Sensor-enhanced Keystroke Dynamics	EER: 4.97%
Roth 2014 [10]	Acoustic	MFCC + K-means	EER: 11%
Zhou 2016[11]	Acoustic	C-SVC + 1-SVM	AA: 92.8%; FRR: 12%; FAR: 11%
WiPass	**WiFi**	**Ricean-*K* + CNN + SVM**	**AA: 92.1%; FRR: 5.9%; FAR: 6.3%**
WiFi based User Identification and Authentication			
	Activity	Algorithm	Performance
WFID 2016 [12]	Gait	Doppler Shift + Radio Scattering + SVM	AA: 93.1% (6 subjects), 91.9% (9 subjects)
WiWho 2016 [13]	Gait	Peak-Valley Detection + DTW + DT	AA: 92% to 80% (2 to 6 users)
Rapid 2017[14]	Gait	CFR + SVM	AA: 82% to 92% (2 to 6 people)
Liu-2018 [15]	Gait	Coherence Time + SVM	AA: 91% (static), 70.6% to 93.6% (mobile)
Shi-2021 [16]	Gait	Neural Network with Auto-Encoder + SVM	Accuracy: 94% / 91% (11 subjects)
NiFi 2019 [17]	Location	Pattern Matching + HMM	True Positive Rate: 90.83% (4 devices)
WiGeR 2016[18]	Gesture	Decomposition + DTW	Accuracy: 97.28% / 91.8% / 95.5% / 83.85%
FingerPass 2020[19]	Gesture	SVM + LSTM	Authentication Accuracy: 91.4%
WiPass	**Keystroke**	**Ricean-*K* + CNN + SVM**	**AA: 92.1%; FRR: 5.9%; FAR: 6.3%**

EER: Equal Error Rate; TPR: True Positive Rate; FPR: False Positive Rate; AA: Authentication Accuracy;
C-SVC: C-Support Vector Classification; SVM: Support Vector Machine; MFCC: Mel Frequency Cepstral Coefficents;
DTW: Dynamic Time Warping; DT: Decision Tree; HMM: Hidden Markov Model; LSTM: Long Short Term Short Memory			

keystroke time) [25] and keystroke forces [26]. But hand-picked features failed to provide a fine-grained view of keystroke dynamics and, thus, recently neural networks have been exploited to improve the authentication performance [27]–[29], as we have done in this article.

B. Wi-Fi-Based User Authentication

Prior Wi-Fi-based user authentication mainly focuses on meter-level body activities, such as gait [12]–[16], since subtle human movements, such as hand gestures and keystrokes, only induce minor channel variations and, thus, are hard to be captured via CSI. Recently, advances in the data-level technologies, such as deep learning, endow us with the ability to enhance the sensing granularity of Wi-Fi via fine-grained data processing, e.g., human presence [17] and hand gestures [19].

However, date-level technologies alone is unable to cope with physical constraints, such as the multipath effect. Therefore, there is a recent trend of exploring signal-level enhancement to improve the Wi-Fi sensing ability to centimeter-level human movements, such as respiration [30] and keystrokes [31]. But previous methods, such as the Fresnel zone model [32], often have strict layout requirements and, thus, could potential offset their scalability and practicality in real world, which drives us to design the Ricean fading based enhancement, which has no layout restrictions and little implementation cost.

C. Positioning WiPass With State-of-the-Art Research

Compared to other keystroke-based authentication approaches using different signals, such as vision [6], [7], sensors [8], [9], and acoustic [10], [11], WiPass constitutes the first Wi-Fi-based authentication approach. It is password-free since we rely on the behaviormetric of keystrokes rather than the password itself for user authentication. Hence, we can use a unified and easy-to-remember phrase for all users rather than a complex password for each user. WiPass is device-free since we exploit the existing Wi-Fi infrastructures for sensing keystrokes instead of mounting any additional devices on the subject. Also, WiPass needs no participation or cooperation of the subject. Moreover, wireless signal is imperceptible for users, easing their privacy concerns.

Compared to other Wi-Fi-based user authentication approaches using different activities, such as presence [17], gait [12]–[16], and gesture [18], [19], WiPass is the first to explore keystroke dynamics for effective user authentication. Desperate that keystrokes is much subtler than previously used activities, WiPass proposes a novel Ricean-*K* based signal enhancement model to effectively improve the sensitivity of channel response to minor human activities and, thus, achieves comparable performance with its state-of-the-art rivals.

III. Ricean Fading Based Signal Enhancement

CSI is a fine-grained physical layer (PHY) information, which characterizes the channel properties of the communication link and provide the amplitude and phase information of each subcarrier. It describes the signal's attenuation factors on each transmission path, such as scattering, multipath fading or shadowing fading, power decay of distance, and other information. In frequency domain, the Wi-Fi channel model can be expressed in terms of CSI as

$$\vec{Y} = \vec{H} \cdot \vec{X} + \vec{N} \tag{1}$$

where $\vec{Y}$ and $\vec{X}$ are the received and transmitted signal vectors, respectively, $\vec{N}$ is the additive white Gaussian noise, and $\vec{H}$ is the channel matrix representing CSI information. Use csitool can extract 30 subcarriers from Intel 5300 network interface card (NIC), for each subcarrier, the channel frequency response (CFR) can be expressed as

$$H_i = L_i + jQ_i = |H_i|e^{j\angle H_i} \tag{2}$$

where i is the subcarrier index, $L_i + jQ_i$ is the CFR value we can obtain from csitool, and $|H_i|$ and $\angle H_i$ are the amplitude and phase of ith subcarrier, respectively.

The Wi-Fi communication in general follows Ricean fading, where the received signal contains two parts, i.e., the LoS components, which describe the direct propagation paths between the transmitter and the receiver, and the NLoS components, which record the diffuse propagation paths between the transmitter and the receiver. In Ricean fading, the Ricean factor is defined as the ratio of the powers of the LoS components to NLoS components, and can be described with the following equation [33]:

$$x(t) = \sqrt{\frac{K\Omega}{K+1}} e^{j(2\pi f_D \cos(\theta_0 t)) + \phi_0} + \sqrt{\frac{\Omega}{K+1}} h(t) \tag{3}$$

where K, Ω, θ_0, and ϕ_0 represent the Ricean factor, the received power, the angle of arrival, and phase of LoS components, respectively. f_D and $h(t)$ denote the maximum Doppler frequency and the NLoS components, respectively.

WiPass maintains stationary after deployment. Therefore, the Doppler frequency (f_D) equals zero. Furthermore, variables, such as Ω, e^{ϕ_0}, and $h(t)$, remain fixed. To this end, (3) can be simplified to

$$x(t) = \sqrt{\frac{K}{K+1}} + \sqrt{\frac{1}{K+1}}. \tag{4}$$

Keystrokes interfere the Wi-Fi transmission and induce dynamic NLoS propagation paths. Therefore, we define H_s and H_d as the static and dynamic channel response as follows, respectively:

$$H_s = \sqrt{\frac{K}{K+1}} + \sqrt{\frac{1}{K+1}} \cdot \rho \tag{5}$$

$$H_d = \sqrt{\frac{1}{K+1}} \cdot (1-\rho) \tag{6}$$

where ρ denotes the ratio of static paths in the NLOS component. So, we divide the received signal into two parts, the static part and the dynamic part, which can be described with the following equation [34]:

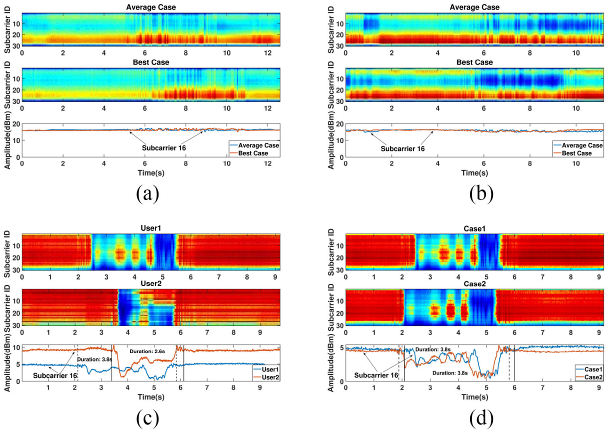
$$|H_{f,\theta}|^2 = |H_s(f)|^2 + |H_d(f)|^2 + 2|H_s(f)||H_d(f)|\cos\theta. \tag{7}$$

Combining (6), we can get

$$\begin{aligned} |H|^2 &= |H_s|^2 + |H_d|^2 + 2|H_s||H_d|\cos\theta \\ &= \frac{K + \rho^2 + 2\sqrt{K}\rho\cos\alpha}{K+1} + \frac{(1-\rho)^2}{K+1} \\ &\quad + \frac{2(1-\rho)\sqrt{K + \rho^2 + 2\sqrt{K}\rho\cos\alpha}}{K+1}\cos\theta \end{aligned}$$

where α is the phase difference of the LOS component to the NLOS, the component in the static part. Considering that the dynamic part contains all the NLoS components, we have $\rho = 0$ and $\alpha = \pi/2$. The aforementioned equation can be further simplified to

$$|f(K)| = |H_s||H_d| = \frac{\sqrt{K}}{K+1} \tag{8}$$



(a) (b)

(c) (d)

Fig. 1. Preliminary experiments verifying the CSI enhancement model. (a) Before enhancement: visualizing keystrokes of user 1. (b) Before enhancement: visualizing keystrokes of user 2. (c) After signal enhancement: comparison between two users. (d) After signal enhancement: comparison between two cases of user 1.

$$|f'(K)| = \frac{1-K}{2\sqrt{K}(1+K)^2}. \tag{9}$$

When $K > 1$, $f(K)$ decreases as K increases. When the moving object is relatively close to the transmitting antenna, as long as the reflector does not reflect the signal away from the moving object, it can ensure that ρ is relatively small. *In other words, by adding an obstacle to reduce the Rice K factor, CSI can be more sensitive to micromotion.*

To this end, we adjust our prototype by deploying a 2.6 (height) × 20 (length) × 0.1 cm (thickness) lead plate, which is large enough for blocking the LoS components according to our experiences, in front of the transit antenna to block the LOS component to enhance the signal distortions caused by micromotions, such as typing. We ask two users (user 1 and user 2) to type the same PIN ten $times$ each under the blocked and unblocked scenarios and compare them with the original cases in Fig. 1. For both users, we notice that the heat map of the averaged blocked case possesses a sharper contrast between the red portion and the blue portion corresponding to the static and dynamic states compared to the unblocked cases, respectively. We further investigate this phenomenon by zooming in at user 1 on the 16th subcarrier where the red and blue lines represent the unblocked and blocked cases, respectively. If focusing on the time period between 6.2 and 10.3 s where the typing happens, we can confirm that the signal distortions are indeed enhanced for the blocked case compared to the unblocked case under the same experimental environment.

After adjusting the prototype for signal enhancement, we restart the preliminary experiments again with the same setting and achieve the following key insight for the design of the WiPass system.

Keystroke dynamics can be leveraged for user authentication: On one hand, when the same person (either user 1 or user 2) strikes the same password twice, there should be strong correlations between the two inputs. On the contrary, even for the same password, user 1 and user 2 have different behavioral dynamics,

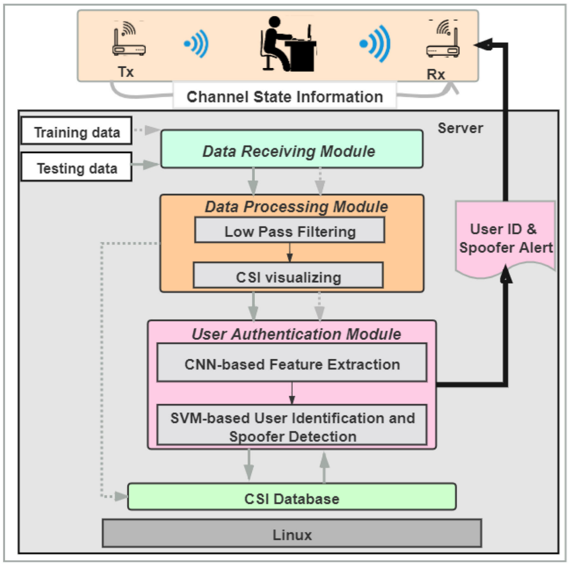


Fig. 2. System architecture of WiPass.

such as the manner and rhythm, in which an individual types on a keyboard. To verify this point, we visualize the enhanced CSI profile corresponding to the scenario in Fig. 1(c) and (d). It is quite interest to see in Fig. 1(c) that different users indeed have different habits of typing. If we focus on the blue part that record the keystrokes, the input time as well as the rhythm is quite different for these two users inputting the same password. We further plot the 16th subcarrier in the lowest image of Fig. 1(c) to show more details. For the same password, user 1 takes 3.8 s while user 2 costs only 2.6 s, which is 46% faster. But if the same user inputs the same password twice, his behavioral habit drives him to input the password in an almost identical manner and rhythm. We take user 1 as an example and look at the 16th subcarrier in the lowest image of Fig. 1(d). The two curves of the amplitude profile also overlap with each other. Also, user 1 takes almost the same time to finish the input, i.e., around 3.8 s.

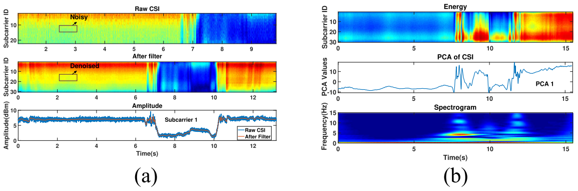
In this section, we have addressed the signal enhancement issue and verified the key idea of leveraging behavioral information for user authentication through preliminary experiments. Based on this foundation, we will present the system design of WiPass in the next section.

IV. System Design

A. System Overview

Fig. 2 presents an overview of the system architecture of WiPass, which mainly consists of two modules, i.e., data processing module and user authentication module. The former aims to remove anomalous data caused by system glitches or environmental interference. The second module targets on user authentication through visualizing the filtered channel data in energy and spectrum domains.

The key idea of WiPass is to leverage the keystroke dynamics, i.e., the manner and rhythm in which an individual types characters on a keyboard, for user authentication. However, previous CSI-based gesture recognition approaches mainly rely on the CSI in the energy domain, which is coarse-grained in describing subtle behavioral information of keystrokes. Therefore, we present a visualization method to utilize the mature CV technologies for extracting behavioral features in both the energy and spectrum domains by mapping the CSI data intro time-series thermal images. Then, we design a hybrid learning model consisting of two components. The first component employs a CNN for feature extraction, whereas the second component uses SVM for user authentication.



(a)

(b)

Fig. 3. (a) Original and filtered CSI time series. (b) Example of visualizing the energy, amplitude, and spectrogram of keystrokes.

B. Data Processing Module

1) Low-Pass Filtering: Raw channel data may contain anomalous samples caused by background noise or surrounding environment. Thus, it should be filtered first, especially considering that the CSI data will be visualized into images for further processing.

The frequency of signal variations due to keystrokes (hands and fingers) lies at the low end of the spectrum, whereas the frequency of the environmental noise usually lies at the high end of the spectrum. To this end, we employ the Butterworth low-pass filter to exclude the background noise. Our preliminary study shows that the frequency of the variations of CSI data caused by keystrokes usually lies within the range of 3–30 Hz. WiPass sets sampling rate $F_s = 1000$ Hz based on our empirical study. Hence, the cut-off frequency of the Butterworth filter is set accordingly

$$\omega_c = \frac{2\pi \cdot f}{F_s} = \frac{2\pi \cdot 30}{1000} = 0.1884(\text{ rand/s}). \tag{10}$$

Fig. 3(a) presents an example to demonstrate the effect of filtering. The unfiltered and filtered CSI waveforms of one complete password input are visualized into two heat-map images [the first two subfigures of Fig. 3(a)], respectively. We further select the subcarrier #1 to have a close look, which is shown in the third subfigure of Fig. 3(a). It is clear that the heat-map image becomes much clearer and sharper after filtering since the noise parts on the original unfiltered image have been removed. Therefore, we can see that the Butterworth filter successfully removes most of the bursty noise from the CSI waveforms while preserving the main sketch comparing to the unfiltered one.

2) CSI Visualizing: The filtered CSI contains important descriptions of keystroke dynamics spread in the time, space, and frequency dimensions. Through data visualization, the cutting-edge data processing methods in CV can be utilized to explore such dynamics in different domains in terms of data features for efficient user authentication. However, the CSI profiles include $N_t \times N_r$ continuous data streams to preserve the keystroke

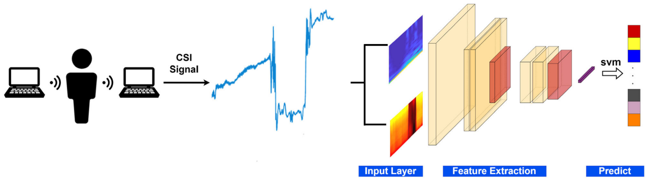


Fig. 4. Architecture of the adopted neural network.

dynamics in both time and space, where N_t and N_r represent the number of transmitting and receiving antennas, respectively. Moreover, for each stream, there exist 30 subcarriers recording the keystrokes in the frequency dimension. Therefore, it is essential to perform dimension reduction to reduce the amount of data being processed by the system to speed up the recognition. In WiPass, we employ the principal component analysis to select the most representative information from CSI to shrink the data volume.

Keystrokes are fine-grained movements that are susceptible to small motions in the surrounding environment. Therefore, it is difficult to characterize the behavioral information using traditional features only in time domains, such as variance or mean of the signal. To tackle this problem, we design a mapping scheme to preserve the behavioral characteristics in both the energy and spectral domains by visualizing the channel response as a time-series image. The energy map contains information about the motion of the hand when typing. Because each person's typing speed is different, the frequency of CSI time-series changes caused by the movement of fingers and hands is also different. Therefore, we convert the time domain signal into frequency domain information through the discrete wavelet transform. The frequency domain information is then converted into a spectrogram. The frequency component on the spectrogram reveals information, such as the frequency of each user typing. Fig. 3(b) illustrates an example of the CSI visualizing method, where the energy map and spectrogram provide rich information of typing in two different dimensions.

C. User Authentication Module

WiPass relies on a traditional two-layer neural network based on a CNN to extract the high-dimensional features of keystroke behavior of each user in order to achieve robust user recognition. As shown in Fig. 4, the network is trained by two channels: spectrum and energy map. The input of the network is a tensor with the shape of (image number) × (image width) × (image height) × (image depth). Before feeding into the network, the energy map and spectrum of one input are first converted into a unified format: 100 × 100. As for each thermal image, its depth means the number of color channels and is set to 3 in WiPass (RGB map). Therefore, the energy map and spectrum map are transformed into 100 × 100 × 3. In addition, we apply the concatenate function to fuse the aforementioned two images into 100 × 100 × 6.

After the input layer reads the features of the two images, the features are fused and then transmitted to the convolution layer.

TABLE II
NETWORK ARCHITECTURE OF CNN MODEL

Layer Type	Filters	Filter size	Output Dimension
Input_1	-	-	100×100×3
Input_2	-	-	100×100×3
Concatenate	-	-	100×100×6
2D-Convolution_1	32	3×3	98×98×32
2D-Convolution_2	32	3×3	96×96×32
2D-Maxpooling	-	2×2	48×48×32
Dropout	-	-	48×48×32
2D-Convolution_3	64	3×3	46×46×64
2D-Convolution_4	64	3×3	44×44×64
2D-Maxpooling	-	2×2	22×22×64
Dropout	-	-	22×22×64
Flatten	-	-	30976
Dense	-	-	256

The convolution layer has a set of learnable filters for convoluting the entire input image and generating various specific types of activation feature maps.

The convolution layer has some key characteristics, such as local perception and parameter sharing, which greatly reduce the number of network parameters, guarantee the sparsity of network, and prevent overfitting. Convolution results are delivered to the next pooling layer, which is essentially a down-sampling operation. The main purpose of the pooling layer is to reduce the dimension of features. Our final pooling result is transformed into a one-dimensional vector and connected to a fully connected layer with 256 hidden layer nodes, and 256 features are obtained. Finally, the high-dimensional features extracted by CNN are used as input of SVM for cross-validation to achieve robust user authentication. The network parameters of WiPass in Table II are set as instructed in [35].

In summary, we explore a CNN for feature extraction since it can better preserve the behavioral information from both time and frequency domains than other traditional classifiers, such as SVM, with manually selected features. We will further verify this idea in extensive real-world experiments.

V. Performance Evaluation

WiPass is device-free in the sense of no extra device needed beyond the existing infrastructure. In our target scenarios, we assume that the existing infrastructure at least consists of a Wi-Fi sender and a receiver, such as our terminal connecting to an access point (AP), so we can leverage the Wi-Fi channel information. Therefore, to simulate such settings, we prototype WiPass on two mini-PCs (2.16GH CPU, 4-GB RAM, and 240-GB SSD) with NIC 5300 as the sender and receiver 60 cm apart from each other. In this way, we can tune the prototype to acquire the desired system parameters for the real-world scenarios, as shown in the prior work [13], [36], [37].

The sampling rate is set to 1000 Hz. The receiver monitors the channel status by the Linux 802.11n CSI tool. We deploy a 2.6 × 20 × 0.1 cm^3 lead plate in front of the transmitting antenna to highlight the signal distortions caused by keystrokes by blocking the LOS components.

We deploy and evaluate WiPass in three real-world environments, whose layouts have been shown in Fig. 5. The 13.62 ×

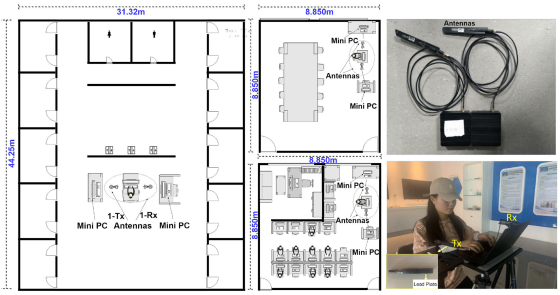


Fig. 5. Our prototype system in three different experimental sites, i.e., lobby, conference room, and laboratory.

44.25 m^2 lobby is relatively empty without much furniture. It mimics the open space setting. The conference room is about 8.85 × 8.85 m^2 with a meeting table and several chairs. Our laboratory is the same size of the conference room but it is quite crowded with students and furniture. To simulate the real-world scenario, there is no restriction on the presence of other people during experiments.

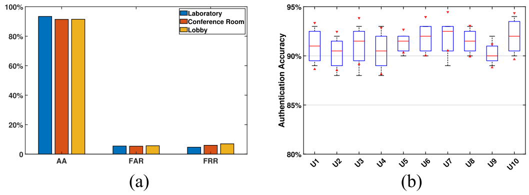
We have recruited 15 participants (6 females) from our university, whose age, weight, and height range from 19 to 27, 38 to 95 kg, and 1.58 to 1.82 m, respectively. Though all of them are familiar with different keyboards, we still ask them to practice on a 104-key Logitech K120 keyboard before the formal experiment. WiPass is a PIN-free authentication system and, therefore, all users are assigned with a unified password to login. Considering the average length of a password is 9.6 characters [38], we set the password of WiPass as "yxx960919," which consists of the initials and birth-date of one female participant. All participants find it easy to remember and easy to input. Later, we will study the impact of passwords on the system performance.

We split 15 participants into two groups, i.e., legitimate users (10) and spoofers (5). Participants in both groups know the password and they are asked to login via the keyboard 30 times at the three different sites. For each person, we set the training size, i.e., the batch size for training, to 20. The dropout ratio and training epoch are set to 0.5 and 30, respectively. We leverage a three-fold cross-validation on the split.

WiPass aims to provide an accurate and reliable user authentication service. Therefore, we define the following evaluation metrics similar to Kong *et al.* [19].

1) Authentication accuracy: the probability that a user can be correctly recognized and authenticated.
2) False accept rate: the probability that a spoofer being authenticated as a legitimate user.
3) False reject rate: the probability that a legitimate user being recognized as a spoofer.

The computation of WiPass is done in our server mounted with two NVIDIA GeForce RTX 2080Ti GPUs, which takes about 58.3 s in training and 2.6 s in testing, respectively.



(a) (b)

Fig. 6. Over evaluation of WiPass. (a) Overall performance. (b) Performance distribution of all legitimate users.

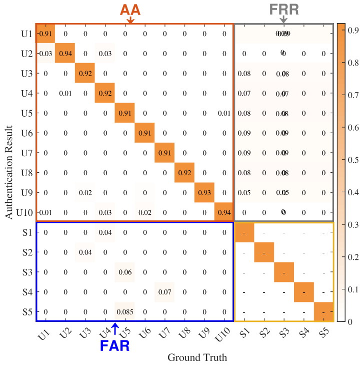


Fig. 7. Averaged confusion matrix over three sites.

A. Overall Performance

Fig. 6(a) shows statistical results over all users in terms of AA, FAR, and FRR in the lobby, conference room, and laboratory, respectively. On average, WiPass achieves 92.1% AA, 5.9% FAR, and 6.3% FRR over the three experimental sites. The results indicate that WiPass is very effective in user authentication and spoofer detection. For instance, Fig. 7 records the confusion matrix over three sites, which has the most complex indoor layout. We can see that WiPass can achieve 92.1% in authentication accuracy.

On one hand, we notice that WiPass performs stably in different real environments. For instance, the gap between the highest and lowest AA, FAR, and FRR, which have been achieved in the laboratory and lobby, is only 2.2%, 0.3%, and 2.3%, respectively. Another interesting observation is that WiPass has a better performance in the laboratory and conference room than the lobby despite the fact that the former two are much smaller and crowded. We think that the direct reason behind this phenomenon is that the multipath effect is caused by the bounce of signal on obstacles, such as furniture and people. In an open area, such effect is fare weaker than in a crowed space.

On the other hand, we find that WiPass is robust to different users. For instance, Fig. 6(b) demonstrates the variation of AA in the cross-validation for all 15 users in the laboratory, where the standard deviation of the averaged AA over 15 users is only 0.0167. We also notice that performance variation is subtle for the same user in the cross-validation, e.g., the maximum standard deviation is 0.0222 of all users in terms of AA. A similar phenomenon has been observed for the other two metrics, i.e., FAR and FRR.

B. How the System Settings Affect the Performance

1) Ricean Distribution: WiPass leverages the Ricean distribution to enhance the subtle signal changes caused by keystrokes. Here, we conduct more experiments by comparing

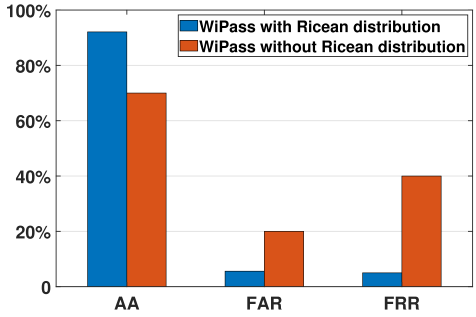


Fig. 8. Impact of Ricean distribution on the performance.

the overall performance with and without the Ricean distribution. Fig. 8 quantifies the performance changes in the laboratory if we removed the lead plate while keeping other settings in terms of AA, FAR, and FRR. It is clear that WiPass suffers a significant performance degeneration with the Ricean distribution, e.g., AA drops from 92.1% to 70.0% while FRR raises from 5% to 40%. In general, it is the Ricean distribution that endows WiPass with the ability of accurate user authentication and spoofer detection. Considering that our methodology is compatible with any similar Wi-Fi-based gesture recognition systems with little implementation method, it constitutes an effective approach for signal enhancements for future research on this topic.

We also notice that FAR and FRR behave differently with the enhancement, i.e., FRR (40%) is twice as large as FAR (20%) before enhancement but both are improved to the same level after enhancement. The reason for this huge difference is because keystroke dynamic only resides in the NLoS components, which are usually much weaker than the LoS components. As a result, even legitimate users could be rejected because their keystroke dynamics tends to be covered by the LoS components, leaving no clear traces for our system to distinguish. After enhancement, these traces would be sharpened and both FRR and FAR have been improved to a similar level.

2) Password: The password is critical to WiPass. We notice that all users think intuitively that our current password (yxx960919) is easy to crack. In contrast, they consider a complex password without a semantic meaning, such as “abc!@&123,” much safer.

Therefore, we ask them to do a back-to-back comparative experiment between those two passwords with the same settings in the laboratory. Surprisingly, the “lousy” password “yxx960919” outperforms the “sound” “abc!@&123” in all three metrics, as shown in Fig. 9(a).

To understand the reason, Fig. 9(b) shows a case study of some user inputting “abc!@&123.” The top subfigure shows the first time this user input the password. Since it is complex, the user paused one time when typing “!,” made one mistake inputting “&,” and then hit the delete key to reinput, leading to a much longer input time compared to the tenth time he inputs without any pause or mistake presented in the middle figure. We focus on subcarrier #1 and highlight where he paused and made the wrong keystroke. Clearly, a complex password is not only hard to remember and input, but also leads to behavioral changes, causing the degeneration in performance.

3) Classifier: Another major component of WiPass is the CNN-based user authentication. We explore the CNN for feature extraction since it can preserve behavioral information in both time and frequency domains. To verify this point, we compare it with two widely used traditional classifiers, i.e., SVM and KNN (K-Nearest Neighbor). Both classifiers utilize the same set of manually picked features, i.e., standards deviation, average absolute error, skewness, kurtosis, entropy, and median. Fig. 9(c) shows the comparison results in terms of AA, where we can see that WiPass achieves the best performance in all three environments. The largest performance gap is found in the lobby, where KNN only achieves 44.44% AA while WiPass yields 89.60% AA, twice as much as the KNN.

4) Keyboard Location: For an ATM or an ACS or even a laptop, the input device is fixed. However, for a desktop, the location of a keyboard can be slightly adjusted. To study the impact of different locations of the keyboard, we conduct experiments in the laboratory where we use the center of the table as the pivot point, and select four other locations 5 cm away from it as the top, bottom, left, and right locations for the keyboard. Fig. 8(d) presents the numerical results, where we can see that the location of the keyboard has an insignificant impact on performance in terms of AA, FAR, and FRR. For instance, the biggest difference in AA is only 2.9%, achieved between the top location and the left location.

To further understand the observation, Fig. 9(e) provides an example of one user inputting the same password at the top and left locations. It seems that the energy maps of both locations share a similar pattern. To further confirm the observation, we focus on the subcarrier #1 in the bottom subfigure, and find that though the location does affect the amplitude profile but the fluctuation pattern of the amplitude remains.

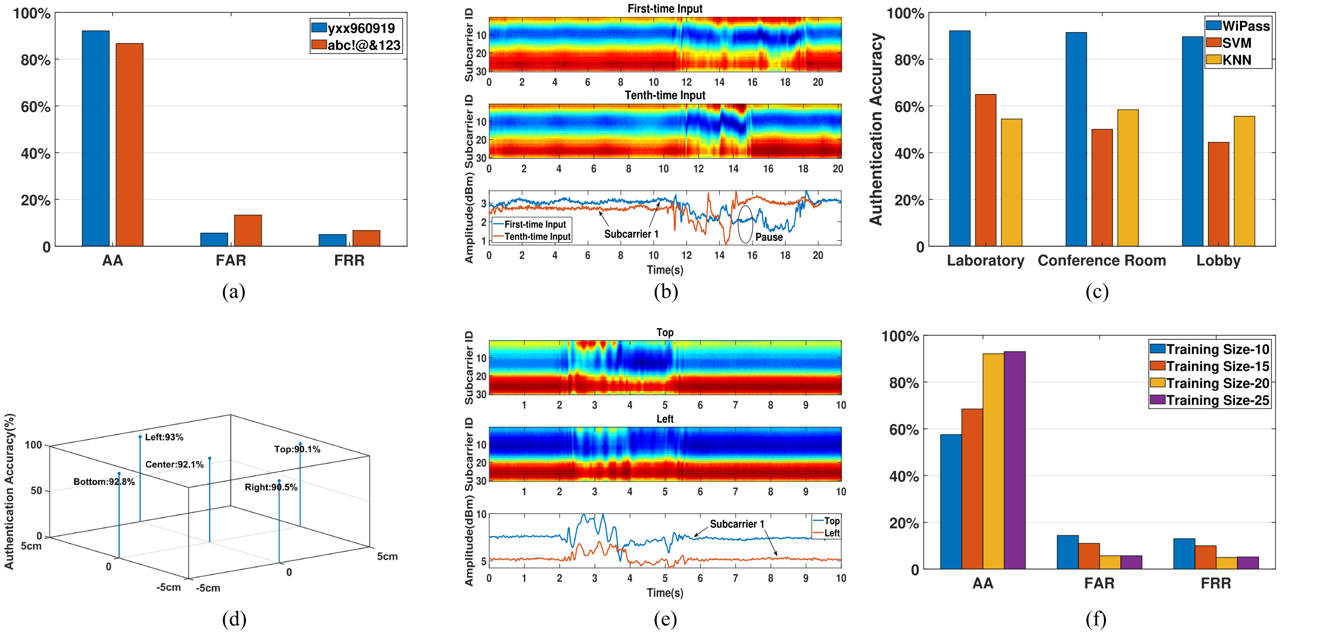
5) Training Size: Fig. 9(f) demonstrates the impact of training size on the performance. In general, WiPass performs better with the increase in the training size. However, the performance improvement in terms of AA, FAR, and FRR becomes stable around 20, which means that the gain brought by a larger training size becomes smaller. Therefore, it is appropriate to choose a training size of 20 to achieve a tradeoff between training cost and performance.

In this part, we have verified the performance of WiPass in real environments. Moreover, we have studied the impact of different system settings on the system. The results not only confirmed robustness and efficiency of WiPass, but also revealed its some interesting features, e.g., a simple password performs better than a complex one.

C. Further Discussions

1) Finger Injury: Finger injury could damage the typing pattern and, thus, deteriorate the performance WiPass. Considering that the index finger is the most often used in keystrokes, we have recruited nine fresh participates (five males and four females, all right-handers) and asked them to perform the following three cases in the lobby.

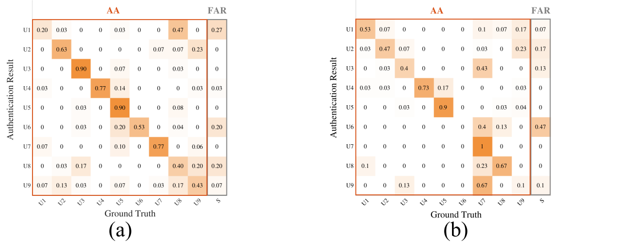
1) Baseline, i.e., no hand injury.



(a) (b) (c)

(d) (e) (f)

Fig. 9. Performance evaluation under different system settings. (a) Impact of different passwords. (b) Case study illustrating why a complex password is not working well for WiPass. (c) Impact of the different classifiers. (d) Impact of different positions of the keyboard. (e) Different locations of the keyboard. (f) Impact of different training sizes



(a) (b)

Fig. 10. (a) Confusion matrix on one-finger injury. (b) Confusion matrix on one-finger injury.

TABLE III
Impact of Human Activities

Scenario	Other Interference	AA	FAR	FRR
Laboratory	No Inference	94.2%	5.5%	4.7%
	Human Typing	93.8%	5.9%	4.7%
	Human Walking	93.0%	5.8%	5.1%
Conference Room	No Inference	93.8%	5.4%	6.0%
	Human Typing	93.4%	5.8%	6.5%
	Human Walking	93.1%	5.7%	6.4%
Lobby	No Inference	91.3%	5.7%	7.0%
	Human Typing	91.2%	5.5%	7.0%
	Human Walking	89.6%	5.8%	7.3%

2) Single-finger case. The participant uses a Band-Aid on the index finger on the right hand.
3) Two-finger case. The participant uses two Band-Aids on the index and middle fingers on the right hand.

Note that all participants are asked to practice phase "yxx960919" for all three cases to make sure they can form a stable typing pattern in the experiments. We use 20 samples of baseline (no injury) as training data and 10 samples of no-injury, single-finger injury, and two-finger injury cases as testing data, respectively. We show their confusion matrices in Fig. 10(a) and (b), where we can see that for the baseline, WiPass can achieve an average 91.3% AA, whereas it drops to 63.7% and 53.3% for single-finger injury and two-finger injury, respectively. It is astonishing to see a single-finger injury could lead to significant change of typing patterns and, thus, deteriorate the performance of WiPass.

2) Human Interference: To understand how human activity interferes WiPass, we recruit nine new volunteers (five males and four females) and ask them to perform the following three cases.

1) Baseline, i.e., no interference.
2) Human typing. Another volunteer is also typing on the keyboard during the experiment.
3) Human walking. Another volunteer is circling around during the experiment.

Note that we employ a 1-m social distance as in real world. Table III summarizes the performance of WiPass under three cases at different sites. On one hand, we notice that the impact of human on WiPass, either typing or walking, is slight. For instance, the average AA over three sites is 93.1% without human interference, whereas the number only drops 0.3% and 1.2% for cases 2 and 3, respectively. This phenomenon is consistent with various prior studies [3], [4], [37], since Wi-Fi signal decays exponentially with the distance and, thus, only leaves insignificant traces on the channel. On the other hand, walking seems to bring more interference than typing. For example, AA reaches 91.2% for typing in the lobby, whereas it drops to 89.6% for walking. This is because typing constrains the location of the human being, whereas walking around could certainly lead to more severe multipath effect.

TABLE IV
INPUT STRINGS WITH DIFFERENCE LENGTHS AND COMPLEXITIES

Input Strings of Different Lengths			
Length	6	11	16
String	YXX919	**YXX919cute!**	YXX0919verycute!
Input Strings of Different Complexities			
Complexity	Easy	Medium	Complex
String	yyy999cccc	**YXX919cute!**	Y9X9c1u!TeX

TABLE V
IMPACT OF LENGTHS ON THE PERFORMANCE

Length of Input Strings	AA	FAR	FRR
Length_6(YXX919) height	91.6%	7.7%	8.1%
Length_11(YXX919cute!)	93.3%	1.7%	6.0%
Length_16(YXX0919verycute!)	90.0%	10.5%	10.3%

TABLE VI
IMPACT OF COMPLEXITY ON THE PERFORMANCE

complexity of Input Strings	AA	FAR	FRR
complexity_easy(yyy999cccc)	91.3%	5.7%	7.0%
complexity_medium(YXX919cute!)	93.3%	1.5%	6.2%
complexity_complex(Y9X9c1u!TeX)	86.7%	13.3%	7.7%

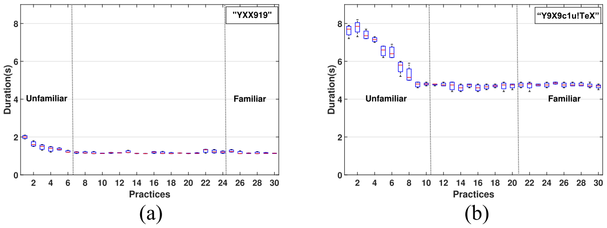


Fig. 11. Duration with different characters. (a) Duration with inputting string "YXX919." (b) Duration with inputting string "Y9X9c1u!TeX."

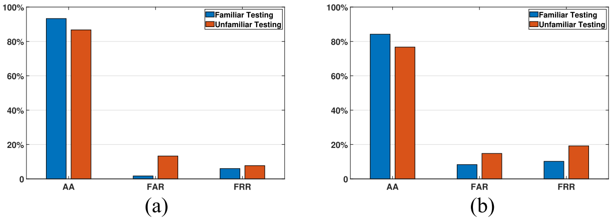


Fig. 12. Results under four kinds of training and testing set with different inputting strings. (a) Performance under different testing set with "YXX919." (b) Performance under different testing set with "Y9X9c1u!TeX."

3) "Password": R. Shay *et al.* conducted a comprehensive survey in the Carnegie Mellon University (CMU) that would elicit truthful information about passwords from 420 CMU computer users, including university students, faculty, and staff, in 2010. They found that the mean of password length is 10.49, whereas the password consists of 5.94 numbers of lowercase letters, 1.54 numbers of uppercase letters (74.2% possibility to be the first character), 2.70 numbers of numbers, and 1.39 numbers of symbols (55.8% possibility to be the first character). Also, the mostly used symbol character is "!." In addition, 43.3% passwords begin with a word/name.

To this end, we design "YXX919cute!," which fits all the aforementioned empirical experiences, as a "pivot" string, based on which we will reshape to study the impact of length and complexity on WiPass. Table IV shows the details of the input strings, where we keep the complexity while varying length and vice versa. We have recruited nine new participants (four male and five female) and conduct new experiments with the settings as in this article. Since our empirical studies have confirmed that WiPass is robust to environments, we only use the lobby for experiments. We conclude the experimental results in Tables V and VI.

Length: On one hand, we notice that the "pivot" string has the best performance, i.e., 93.3% AA, 1.7% FAR, and 6.0% FRR. It fits well with our real-world experiences with passwords and, thus, is easily accepted by the participants. On the other hand, we notice that though "YXX0919verycute!" is also easy to understand its performance in terms of FAR and FRR degenerates much from the "pivot" string. The participants have trouble fluently typing this string and, thus, has hesitations between letters. Such hesitation is rather random and damages individual's keystroke dynamics.

Complexity: It is quite interesting that the "pivot" still has the best performance among three strings with different complexities. We think an oversimplified string, such as "yyy999ccc," requires much less finger movements and, thus, damages the typing pattern. But a quite complex string, such as "Y9X9c1u!TeX," really troubles all the participants and makes them think hard during typing. The thinking will certain induce pauses between letters and, thus, add noises to the keystroke dynamics. As a result, it has the worst performance among all the five strings, i.e., 86.7% AA, 13.3% FAR, and 7.7% FRR.

4) Familiarity Issue: Our prior experiments are based on the fact that participants are familiar with the input strings after 30 times' practice. But how much impact does the familiarity issue have on the authentication results, especially for the complex ones?

To this end, we focus on two extreme cases, i.e., the simplest string "YXX919" and the most complex string "Y9X9c1u!TeX," and study the impact of familiarity on their performance. More specifically, we concentrate on the practice process itself instead of the experiments after practice. Fig. 11 records the distributions of input duration of "YXX919" and "Y9X9c1u!TeX." We notice that it takes the participants 5.8 and 10.1 practices on average for "YXX919" and "Y9X9c1u!TeX" before the input duration becomes stable, respectively. It is quite naturally since the latter string is much difficult to remember and type. Assuming the reliable keystroke dynamics lies in the latter part of the practice, we use the unfamiliar parts (6 for "YXX919" and 10 for "Y9X9c1u!TeX") as testing to study the impact of familiarity, and conclude the results in Fig. 12.

On one hand, we notice that familiarity does affect the performance of WiPass for both cases. For instance, WiPass still achieves 86.7% AA, 13.3% FAR, and 7.7% FRR with the unfamiliar practices. On the other hand, we observe that FAR drops faster than FRR for "YXX919," whereas the observation is opposite for "Y9X9c1u!TeX." We think it is because a simple string possesses less distinguishable typing features and, thus, participants tend to form a similar pattern, leading to more false accepts. On the contrary, a complex string usually leads to diversified type patterns among different participants, resulting in more false rejects.

VI. Conclusion

This article introduced WiPass, a device-free and PIN-free user authentication system leveraging commodity Wi-Fi. The key idea was that we explored the way the PIN was typed (behavioral information) other than the PIN itself to recognize the user. To characterize the behavioral features hidden in the human subtle micromotions, such as keystrokes, we built a signal enhancement model using Ricean distribution and a hybrid learning model for user authentication, which consists of two parts, i.e., CNN-based feature extraction and SVM-based classification. The former relied on visualizing the channel responses into time-series images to learn the behavioral features in energy and spectrum domains, whereas the latter exploited such behavioral features for user authentication. We realized WiPass on commodity Wi-Fi devices and verified its performance in three real environments. For the future work, we would focus on the robustness issue of WiPass under various factors, such as different attack models (such as impersonation attack and substitution attack) and different scenarios (such as ATMs and mobiles).

Acknowledgment

The authors would like to thank the anonymous reviewers for their time and effort.

References

[1] A. Hanamsagar, S. S. Woo, C. Kanich, and J. Mirkovic, "Leveraging semantic transformation to investigate password habits and their causes," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2018, Art. no. 570.

[2] Data-breach, "Data-breach," 2020. [Online]. Available: https://www.idagent.com/blog/10-facts-about-passwords-that-you-need-to-see-now/

[3] Y. Gu, J. Zhan, Y. Ji, J. Li, F. Ren, and S. Gao, "MoSense: An RF-based motion detection system via off-the-shelf Wi-Fi devices," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2326–2341, Dec. 2017.

[4] Y. Gu, X. Zhang, Z. Liu, and F. Ren, "BeSense: Leveraging WiFi channel data and computational intelligence for behavior analysis," *IEEE Comput. Intell. Mag.*, vol. 14, no. 4, pp. 31–41, Nov. 2019.

[5] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless respiration monitoring via off-the-shelf WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2466–2479, Oct. 2016.

[6] J. Roth, X. Liu, and D. Metaxas, "On continuous user authentication via typing behavior," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4611–4624, Oct. 2014.

[7] B. Ayotte, J. Huang, M. K. Banavar, D. Hou, and S. Schuckers, "Fast continuous user authentication using distance metric fusion of free-text keystroke data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 2380–2388.

[8] H. Lv and W.-Y. Wang, "Biologic verification based on pressure sensor keyboards and classifier fusion techniques," *IEEE Trans. Consum. Electron.*, vol. 52, no. 3, pp. 1057–1063, Aug. 2006.

[9] C. Giuffrida, K. Majdanik, M. Conti, and H. Bos, "I sensed it was you: Authenticating mobile users with sensor-enhanced keystroke dynamics," in *Proc. Int. Conf. Detection Intrusions Malware, Vulnerability Assessment*, 2014, pp. 92–111.

[10] J. Roth, X. Liu, A. Ross, and D. Metaxas, "Investigating the discriminative power of keystroke sound," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 2, pp. 333–345, Feb. 2015.

[11] Q. Zhou, Y. Yang, F. Hong, Y. Feng, and Z. Guo, "User identification and authentication using keystroke dynamics with acoustic signal," in *Proc. 12th Int. Conf. Mobile Ad-Hoc Sensor Netw.*, 2016, pp. 445–449.

[12] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "WFID: Passive device-free human identification using WiFi signal," in *Proc. 13th Int. Conf. Mobile Ubiquitous Syst.: Comput., Netw., Serv.*, 2016, pp. 47–56.

[13] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, 2016, pp. 1–12.

[14] Y. Chen, W. Dong, Y. Gao, X. Liu, and T. Gu, "Rapid: A multimodal and device-free approach using noise estimation for robust person identification," *Proc. ACM Interactive, Mobile, Wearable, Ubiquitous Technol.*, vol. 1, no. 3, pp. 1–27, 2017.

[15] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 251–264, Feb. 2018.

[16] C. Shi, J. Liu, H. Liu, and Y. Chen, "WiFi-enabled user authentication through deep learning in daily activities," *ACM Trans. Internet Things*, vol. 2, no. 2, pp. 1–25, 2021.

[17] L. Cheng and J. Wang, "Walls have no ears: A non-intrusive WiFi-based user identification system for mobile devices," *IEEE/ACM Trans. Netw.*, vol. 27, no. 1, pp. 245–257, Feb. 2019.

[18] M. A. A. Al-qaness and F. Li, "WiGeR: WiFi-based gesture recognition system," *ISPRS Int. J. Geo- Inf.*, vol. 5, no. 6, 2016, Art. no. 92.

[19] H. Kong, L. Lu, J. Yu, Y. Chen, L. Kong, and M. Li, "FingerPass: Finger gesture-based continuous user authentication for smart homes using commodity WiFi," in *Proc. 20th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2019, pp. 201–210.

[20] B. Mag., "First of the sender," 2021. [Online]. Available: http://www.baselinemag.com/c/a/Security/The-Rhythm-of-Identity-Management

[21] A. A. N. Buker, G. Roffo, and A. Vinciarelli, "Type like a man! Inferring gender from keystroke dynamics in live-chats," *IEEE Intell. Syst.*, vol. 34, no. 6, pp. 53–59, Nov./Dec. 2019.

[22] A. Ogihara, H. Matsumura, and A. Shiozaki, "Biometric verification using keystroke motion and key press timing for ATM user authentication," in *Proc. Int. Symp. Intell. Signal Process. Commun.*, 2006, pp. 223–226.

[23] M. Abuhamad, A. Abusnaina, D. Nyang, and D. Mohaisen, "Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 65–84, Jan. 2021.

[24] R. Joyce and G. Gupta, "Identity authorization based on keystroke latencies," *Commun. ACM*, vol. 33, no. 2, pp. 168–176, 1990.

[25] F. Monrose and A. Rubin, "Authentication via keystroke dynamics," in *Proc. 4th ACM Conf. Comput. Commun. Secur.*, 1997, pp. 48–56.

[26] K. Kotani and K. Horii, "Evaluation on a keystroke authentication system by keying force incorporated with temporal characteristics of keystroke dynamics," *Behav. Inf. Technol.*, vol. 24, no. 4, pp. 289–302, 2005.

[27] D. T. Lin, "Computer-access authentication with neural network based keystroke identity verification," in *Proc. Int. Conf. Neural Netw.*, 1997, pp. 174–178.

[28] L. Sun, Y. Wang, B. Cao, P. S. Yu, W. Srisa-An, and A. D. Leow, "Sequential keystroke behavioral biometrics for mobile user identification via multi-view deep learning," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2017, pp. 228–240.

[29] Y. Gu *et al.*, "WiONE: One-shot learning for environment-robust device-free user authentication via commodity Wi-Fi in man-machine system," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 3, pp. 630–642, Jun. 2021.

[30] Y. Gu, C. Zhang, Y. Wang, Z. Liu, Y. Ji, and J. Li, "A contactless and fine-grained sleep monitoring system leveraging WiFi channel response," in *Proc. IEEE Int. Conf. Commun.*, 2019, pp. 1–5.

[31] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Recognizing keystrokes using WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1175–1190, May 2017.

[32] D. Zhang, F. Zhang, D. Wu, J. Xiong, and K. Niu, "Fresnel zone based theories for contactless sensing," *Contactless Hum. Activity Anal.*, vol. 200, pp. 145–164, 2021.

[33] C. Tepedelenlioglu, A. Abdi, and G. B. Giannakis, "The Ricean K factor: Estimation and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 799–810, Jul. 2003.

[34] H. Wang *et al.*, "Human respiration detection with commodity WiFi devices: Do user location and body orientation matter?," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 25–36.

[35] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[36] M. Li *et al.*, "When CSI meets public WiFi: Inferring your mobile phone password via WiFi signals," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 1068–1079.

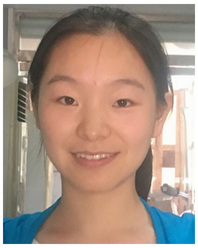
[37] Y. Gu, F. Ren, and J. Li, "PAWS: Passive human activity recognition based on WiFi ambient signals," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 796–805, Oct. 2016.

[38] J. Lampe, "Beyond password length and complexity," 2014. [Online]. Available: https://resources.infosecinstitute.com/beyond-password-length-complexity

Yu Gu (Senior Member, IEEE) received the B.E. degree from the Special Classes for the Gifted Young, University of Science and Technology of China, Hefei, China, in 2004, and the D.E. degree from the University of Science and Technology of China, in 2010, both in computer science.

In 2006, he was an Intern with Microsoft Research Asia, Beijing, China, for seven months. From 2007 to 2008, he was a Visiting Scholar with the University of Tsukuba, Tsukuba, Japan. From 2010 to 2012, he was a JSPS Research Fellow with the National Institute of Informatics, Tokyo, Japan. He is currently a Professor with the School of Computer and Information, Hefei University of Technology, Hefei. His current research interests include pervasive computing and affective computing.

Dr. Gu was the recipient of the IEEE Scalcom2009 Excellent Paper Award and NLP-KE2017 Best Paper Award. He is a Member of ACM.



Yantong Wang received the B.E. degree in computer science and technology from Shanghai Normal University, Shanghai, China, in 2016, and the master's degree in computer science and technology in 2020 from the Hefei University of Technology, Hefei, China, where she is currently working toward the Ph.D. degree.

Her research interests include affective computing and wireless sensing.

Meng Wang received the bachelor's degree in Internet of Things engineering in 2020 from the Hefei University of Technology, Hefei, China, where she is currently working toward the Ph.D. degree.

Her current research interests include affective computing, wireless sensing, and machine learning.

Zulie Pan was born in China, in 1976. He received the B.E. degree in computer and applications from the Information Engineering University, Zhengzhou, China, in 1997, the master's degree in communication science from the Electronic Engineering Institute, Hefei, China, in 2004, and the Ph.D. degree in signal and information processing from the Electronic Engineering Institute, Hefei, China, in 2009.

He is currently a Professor with the National University of Defense Technology, Hefei. His main research interests include network security and computer science.

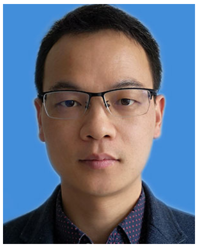
Zhihao Hu received the B.E. degree in network engineering from the National University of Defense Technology, Hefei, China, in 2019, where he is currently working toward the master's degree in network security.

His main research interests include network security and fuzzing.

Zhi Liu (Senior Member, IEEE) received the B.E. degree from the University of Science and Technology of China, Hefei, China, in 2019 and the Ph.D. degree in informatics from the National Institute of Informatics, Tokyo, Japan, in 2014.

He is currently an Associate Professor with The University of Electro-Communications, Tokyo. His research interests include video network transmission and mobile edge computing.

Dr. Liu is currently an Editorial Board Member of Springer *Wireless Networks* and IEEE Open Journal of the Computer Society.



Fan Shi was born in China, in 1983. He received the B.Sc. (hons.) degree in network engineering and the M.Sc. degree in information security from the National University of Defense Technology, Hefei, China, in 2004 and 2007, respectively.

He is currently an Associate Professor with the National University of Defense Technology. His research interests include information security and network security situation awareness.

Mianxiong Dong (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer science and engineering from The University of Aizu, Aizuwakamatsu, Japan, in 2006, 2008, and 2013, respectively.

He is the youngest ever Vice President and Professor with the Muroran Institute of Technology, Muroran, Japan. He was a JSPS Research Fellow with the School of Computer Science and Engineering, The University of Aizu and was a Visiting Scholar with BBCR Group, University of Waterloo, Waterloo, ON, Canada, supported by JSPS Excellent Young Researcher Overseas Visit Program from April 2010 to August 2011.

Dr. Dong was selected as a Foreigner Research Fellow (a total of three recipients all over Japan) by NEC C&C Foundation in 2011. He was the recipient of the IEEE TCSC Early Career Award 2016, IEEE SCSTC Outstanding Young Researcher Award 2017, the 12th IEEE ComSoc Asia-Pacific Young Researcher Award 2017, Funai Research Award 2018, and NISTEP Researcher 2018 (one of only 11 people in Japan) in recognition of significant contributions in science and technology. He is Clarivate Analytics 2019 Highly Cited Researcher (Web of Science).