Your Breath Doesn't Lie: Multi-user Authentication by Sensing Respiration Using mmWave Radar

Yao Wang[1], Tao Gu[2], Tom H. Luan[1], and Yong Yu[3]

[1]School of Cyber Engineering, Xidian University, China
[2]School of Computing, Macquarie University, Australia
[3]School of Cyber Security, Xi'an University of Posts and Telecommunications, China
{wangyao, tom.luan}@xidian.edu.cn, tao.gu@mq.edu.au, yongyuxy@163.com

2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON) | 978-1-6654-8643-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/SECON55815.2022.9918606

***Abstract*—User authentication is critical to privacy preservation. Most of the existing works focus on single-user authentication, which may not work efficiently and practically in multi-user scenarios. To this end, we present a *M*ulti-user *Auth*entication system (*M-Auth*) that employs a single COTS mmWave radar to capture the user's unique breathing pattern. It exploits the phenomenon that radio frequency (RF) signals are affected by chest displacements due to breathing. We specifically design an auxiliary rotating gadget to dynamically adjust radar orientation, making it more effective in capturing respiration signals from multiple users. To profile individual components from the entangled RF signals, we leverage mmWave's high directivity to locate each user and separately focus on reflections from different positions. We propose a signal energy comparison method to eliminate the irrelevant body movements for preserving fine-grained respiration traits. Afterward, we develop a feature selection pipeline to elicit the most informative features and train a machine learning-based classifier to identify each user. *M-Auth* is practical due to its non-contact and passive nature, and it is secure as respiration is unique and difficult-to-forge. Extensive experiments involving 37 participants demonstrate that *M-Auth* is effective in verifying legitimate users and thwarting spoofing attacks, with an authentication accuracy of over 96% and an attack detection rate of over 95%.**

***Index Terms*—Respiration, Authentication, mmWave Sensing**

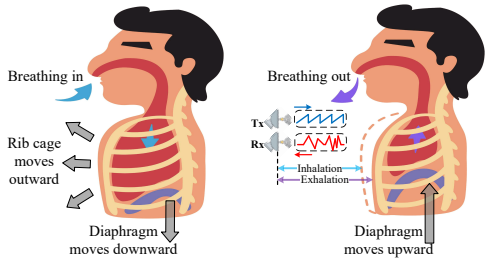
I. Introduction

User authentication mechanisms have evolved from complex passwords to biometrics (e.g., fingerprint, iris, and voice). One important thing to realize is that biometric scans may still be thwarted. Attackers have defeated biometric security with impressions of fingerprints, a contact lens placed over a photo of an iris, or recording someone's voice and playing it back for a voice recognition system. The problem with these methods is that they only provide one-time authentication. In many application scenarios such as banking and home computing, continuous authentication is often required to validate user identity constantly throughout the entire session.

Existing solutions typically exploit behavioral biometrics such as gait patterns [1] and keystroke dynamics [2] to perform continuous authentication. However, these solutions require users to actively engage with the authentication process, such as keep walking within a specific range and typing on the keyboard. To avoid such laborious operations, non-contact continuous authentications have been proposed by using wireless sensing (e.g., WiFi and RFID) and spontaneous physiological biometrics (e.g., respiration and heartbeat). For example, *Cardia Scan* [3] and *BreathID* [4] employs a continuous-wave Doppler radar and WiFi signals to capture heartbeat and respiration motions for continuous authentication, respectively. They have the advantage of freeing users from getting involved in authentication, the following significant limitations have yet to be resolved. For one thing, their restricted working range means that it will be inaccurate for far-field users, which is inappropriate in larger spaces. For another thing, they just perform single-user authentication, ignoring the broader applications of multi-user settings such as smart homes and workplaces, where more than one person is usually present.

Recent efforts have been made to enable multi-user authentication. For instance, Kong *et al.* reuse WiFi signals to capture several predefined activities from different users and employ time-of-arrival to distinguish the multiple components [5]. Although allowing multi-user authentication, the proposed system requires at least $0.8m$ spacing between users to achieve an acceptable accuracy, which poses challenges when users are in close proximity such as standing shoulder to shoulder or sitting abreast. Besides, it still has the constraint of requesting users to complete specified tasks.

Design Objectives. In this paper, we present *M-Auth*, a continuous multi-user authentication system by sensing respiratory motions with a single COTS mmWave radar. Before being authenticated, users enroll in the system to create profiles, and an incoming respiration signal is compared to the stored profiles to identify whether the signal is from a genuine user or an attacker. Specifically, the highlights of our work are as follows: **1) Ubiquitous and trustworthy.** All living people have to breathe and breathing motions are difficult to forge, while gait or hand-based approaches are hardly applicable to individuals with hand or foot disabilities and may readily be acquired by a peripheral camera for imitation attacks [6]. **2) Zero separation spacing.** Our system enables multi-user authentication even when users have zero distance between them, whereas existing RF-based solutions typically need at least a separation of 0.8-1m [1], [5], which obviates scenarios such as when users sit close to each other. **3) Non-contact and passive.** We employ RF waves to remotely perceive the unique naturally-occurring respiration motion, which does not involve device touch or physical exertion.



(a) Inhale Motion (b) Exhale Motion

Fig. 1. Respiratory motion cycle and sensing rationale.

Technical Challenges. Recent studies have used mmWave sensing to estimate respiratory rate by monitoring signal peaks in a target user's reflections [7], [8]. However, unlike these coarse-grained measurements, authentication involves identifying subtle differences between users' respiration signals. This requires us to solve the following challenges: 1) The angle-of-arrival (AoA) of a signal sensitively affects its signal-to-noise ratio (SNR), larger AoA might weaken the signal strength, leading to inaccurate authentication. Existing beamforming-related techniques such as phased array and beam steering typically improve spectral efficiencies in a particular direction [9], which present challenges in our scenario since we assume users are mobile and their AoAs are unpredictable. On account of this, we design a rotating device that mechanically controls the radar to dynamically adjust its orientation according to the users' positions, ensuring effective signal capture. 2) Hand and limb movements may overwhelm the small chest displacements due to respiration. We present a comparison scheme to remove such interference by calculating the signal energy for a specific time window. 3) Accurate authentication depends on highly recognizable features. For this purpose, we develop a feature selection pipeline that combines wavelet packet decomposition (WPD) and recursive feature elimination (RFE) methods to select the most representative features from the respiration signal.

Applications. In today's IoT-rich environments, it is convenient when multi-user authentication functionality is supported. *M-Auth* can be deployed at the building entrance or in the multi-person office to verify personnel for access control. Smart homes may use *M-Auth* to associate indoor persons with their identities for security surveillance, parent control, as well as certain personalized services such as heating, ventilation, and air conditioning (HVAC) applications. It can also be harnessed to increase the continuous authentication capability for traditional one-time confirmation mechanisms. In summary, this paper makes the following contributions:

1) We develop a continuous authentication system for multiple users based on unique and non-volitional breathing patterns. This system is secure, passive, and contactless, offering a practical solution for various applications.

2) We design an auxiliary rotating device to enable a single mmWave radar to dynamically adjust its orientation for effective sensing. This design can potentially transform existing statically deployed sensing solutions into adaptive, cost-effective, and reliable coverage.

3) We demonstrate the effectiveness of *M-Auth* through extensive experiments with 37 participants. Results show that *M-Auth* achieves an accuracy of more than 96%. We conduct a range of experimental studies to evaluate the performance under multiple scenes and various attacks.

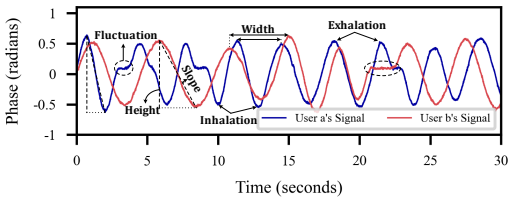


Fig. 2. Variations of phase due to respiration.

II. Preliminaries

A. Respiratory Biometrics

This paper exploits respiratory motion as the biometric factor to perform user authentication. Typically, one respiratory cycle includes two phases–inhalation and exhalation. In the process of inhalation, as shown in Fig. 1(a), the intercostal muscles and the abdominal muscles contract to pull out the ribs, and the diaphragm moves downward to be flat, resulting in the expansion of the chest cavity. While during exhalation, as shown in Fig. 1(b), the muscles relax and the diaphragm moves upward to return to its resting position, leading to a decrease in the size of the chest cavity. The two phases vary from person to person due to human physiological structure, such as different lung volumes and various acceleration of chest moving dynamics. In addition, since respiratory motion is inherently related to physiological activities, it is more difficult to forge than traditional biometrics (e.g., face and voice). Therefore, we adopt respiratory motion as the unique biometric modality for user authentication.

B. Feasibility Study

It is hypothesized that a frequency modulated continuous wave (FMCW) signal can be leveraged to sense the chest movements due to respiration, and it is capable of distinguishing the minute differences between individuals. As illustrated in Fig. 1(b), the Tx transmits periodic sawtooth waves to the user. When breathing, the chest fluctuations modulate the incident signal, and the Rx will sense the chest-reflected signal. Since signal phase and distance are linearly dependent, chest displacements can be tracked by calculating the phase changes between consecutive measurements. Specifically, variations of the distance $d(t)$ can be calculated as follows [10]:

$$d(t) = \frac{\lambda}{4\pi}\phi(t), \tag{1}$$

where λ and $\phi(t)$ are the wavelength and phase, respectively. It is observed that short-wavelength signals are more sensitive to distance variations. In this work, we use a $4mm$ wavelength mmWave, Δd can achieve as high as $1mm$ when the phase change $\Delta\phi$ is π, which is competent to detect the small chest displacements produced by respiration.

We further investigate how respiration diversity influences the captured signals. Two participants are asked to sit facing the device and breathe normally. Fig. 2 shows their respiration

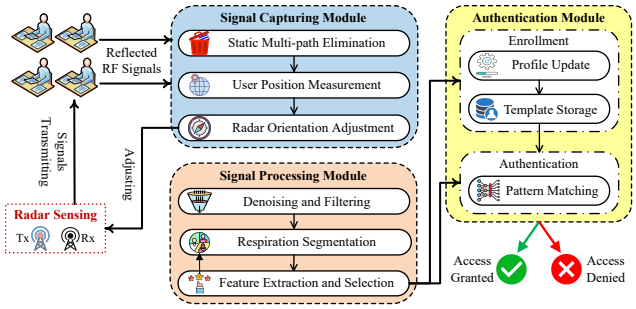
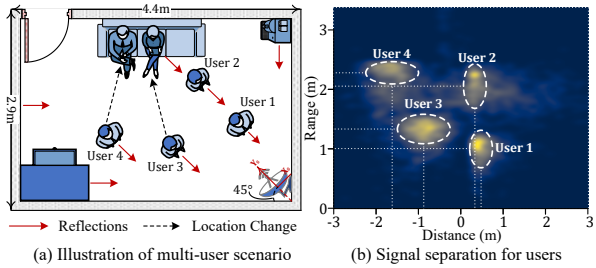


Fig. 3. System overview of *M-Auth*.



(a) Illustration of multi-user scenario (b) Signal separation for users

Fig. 4. Signal separation for multiple users and user location measurement.

waveforms. Intuitively, it is observed that the signals are significantly different between the two users. Morphological characteristics such as pulse height, width, slope, and fluctuations apparently change from User a to User b. These dissimilarities are mainly caused by individual differences in the strength of intercostal muscles and lung volume. This study demonstrates that mmWave can capture the minute differences in respiratory motions, motivating us to use such user-specific traits to perform authentication.

C. Threat Model

We assume that the end device is secure, attacks such as tampering matching mechanism and stealing biometric templates are orthogonal to this work. Although respiration motion is complex and might be more secure than other authentication modalities such as password and fingerprint, to verify its reliability, we consider the following attacks:

Blind Attack. An adversary is unclear about the genuine user's breathing patterns, e.g., rate, depth, and rhythm changes. During the attack, the adversary performs random respiration motions to *M-Auth*, hoping to produce similar impacts on the system as the genuine user does.

Impersonation Attack. An adversary can observe the legitimate user's respiration motions by shoulder surfing or video recordings. The adversary tries to mimic the user's breathing patterns according to his/her own understanding.

Replay Attack. More advanced than the former two attacks, we assume 1) the adversary knows the authentication principle and places a hidden mmWave sensor in a nearby location to record the legitimate user's body-reflected signals. 2) The adversary can eavesdrop on the internal communication and inject the recorded signal to spoof the system.

III. System Overview

Fig. 3 presents the workflow of *M-Auth*, which consists of the following modules:

Signal Capturing Module. Since beamforming techniques generally reinforce the signals in a specific direction, it is not suitable for improving signal quality when users are movable. To effectively capture echo signals from multiple users, we first eliminate the reflections from static objects (e.g., walls) and measure the users' positions. Then we implement a clustering algorithm to estimate the central position of the users. Lastly, we develop a mechanical rotating device that assists our radar in dynamically adjusting its orientation to face toward the centroid for reliable sensing.

Signal Processing Module. Once the direction is determined, *M-Auth* starts to authenticate the users who are in the range of the radar. This module eliminates noises from the captured signal by combining a band-pass filter and an adaptive filter. To remove reflections from irrelevant body motions, we introduce a signal comparison scheme by calculating the signal energy for a specific time window. Subsequently, the system segments the respiration signal by extremum analysis. At last, we use the WPD and RFE techniques to select the features that are most associated with respiration.

Authentication Module. After feature selection, *M-Auth* labels the corresponding features and saves them to construct the biometric templates. We build a machine learning-based classifier by using the stored features to determine whether an unknown visitor is a legitimate user or an attacker. Besides, respiration patterns might change a lot due to mood swings and energetic exercises, our system also provides template updating to adapt to the changes under such cases.

IV. Signal Capturing Module

A. Signal Separation for Different Users

We consider a multi-person scenario as illustrated in Fig. 4(a), where a mmWave radar is deployed in the corner of the room that has four persons and several furniture. To identify human-reflected signals from those reflected off other static objects and to separate individual signals from the multiple users in the environment, we leverage the intrinsic property of FMCW, which enables separating the reflections from different objects. In what follows, we go through the specifics of signal capture and isolation from multiple users.

Static Multi-path Elimination. For an object at a distance d from the radar, the radar mixes the TX and RX chirps to generate an intermediate frequency (IF) signal. Given the slope of the chirp S, the frequency of the IF signal is [10]:

$$f = S \cdot \tau = \frac{B}{T_c} \cdot \frac{2d}{c} = \frac{2Bd}{cT_c}, \tag{2}$$

where τ, B, T_c, and c are the time delay of the RX chirp, frequency bandwidth, chirp duration, and speed of light, respectively. As we can observe from Eq. (2), since the distance of static reflectors (e.g., furniture) to the radar is constant over time, the induced frequency shift does not change over time. Consequently, we can get rid of those time-invariant multi-path reflections by subtracting consecutive time measurements.

User Presence Detection. When a user appears in the radar's field-of-view (FoV), our system receives the reflected signal from the user. Since body movements cause changes in d, consequently triggering strong responses in the IF signal according to Eq. (2). We use this phenomenon to detect the presence of users in the environment, and further estimate their positions in the next step.

User Position Measurement. Merely using range information d is insufficient to distinguish between multiple users since they are likely to have similar distances to the radar but be in different directions. Therefore, we introduce another horizontal distance parameter to determine the user position. For instance, the horizontal distance from User 1 to the radar is calculated as $d_1 \sin\theta_1$, where θ_1 represents the angle of arrival (AoA) that is measured as follows [10]:

$$\theta_1 = \sin^{-1}(\frac{\lambda\Delta\phi_1}{2\pi l}), \tag{3}$$

where λ, $\Delta\phi_1$, and l are the signal wavelength, phase change, and spacing between RX antennas, respectively. Accordingly, the $n'th$ user's position is expressed as $P_n(d_n \sin\theta_n, d_n)$.

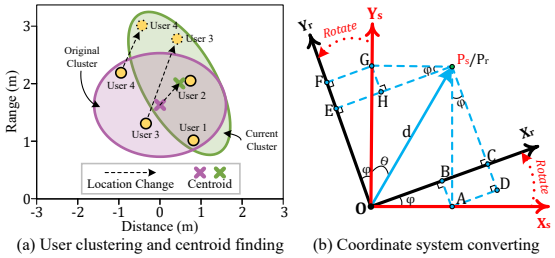
Multi-user Signal Separation. For multiple objects that are present in an environment, each RX chirp is separated by a different amount of time delay which is proportional to the distance from the system to the object. In this case, a Fourier transform is used to process the IF signal consisting of multiple tones, resulting in a frequency spectrum with discrete peaks for each tone, each peak indicating the presence of an object at a certain distance. Further, on the basis of Fourier transform theory, frequency components can be separated as long as their frequency difference Δf is more than $\frac{1}{T_c}Hz$ [10], where T_c is the chirp duration. By using Eq. (2), the relationship is represented as:

$$\Delta f = \frac{2B\Delta d}{cT_c} > \frac{1}{T_c} \Rightarrow \Delta d > \frac{c}{2B}. \tag{4}$$

In this work, the radar provides a $4GHz$ bandwidth, such that the range resolution Δd is calculated as $\frac{c}{2B} = \frac{3\times10^8}{2\times4\times10^9} = 3.75cm$. This means that if objects are at least $3.75cm$ apart, their received chirps can be identified separately. In our context, we primarily focus on user's chest movements caused by respiration, even when users are standing shoulder to shoulder (i.e., zero separation distance between them), our system is still capable of distinguishing the RX chirps from them since their chest positions are separated by arms (the spacing is typically more than $3.75cm$). As shown in Fig. 4(b), the reflections from multiple users are separated into areas, allowing us to further analyze their signals individually.

B. Dynamic Radar Orientation Adjustment

Design Motivation. In this work, we use a MIMO radar that supports an FoV of 120°. Intuitively, it is not necessary to adjust radar orientation since a radar statically positioned in the corner of the room with configured beamforming can cover all the users. However, radar's phase change $\Delta\phi$ is sensitive to AoA and the estimation of $\Delta\phi$ degrades as AoA increases [11]. In other words, when users draw closer to the border, the measurement of respiration motions becomes more error-prone. Our experiment in Section VIII-C also confirms that when AoA changes from 0° to 60°, the authentication accuracy decreases by approximately 10%.



(a) User clustering and centroid finding (b) Coordinate system converting

Fig. 5. Centroid estimation and location mapping.

Signal processing techniques such as adaptive beamforming, beam steering, and beam switching might be applied to address this issue [9]. However, these methods improve the signal SNR in a specific direction, i.e., they assume that the object's position is fixed. In our scenario, users are mobile and their AoAs are uncertain, signal processing-based methods are consequently not applicable. To this end, we propose an approach that physically adjusts the radar's orientation in real-time depending on the user's location as follows.

Step 1 - Radar Direction Calibration. Initially, we calibrate the sensor coordinate system to align with the room coordinate system, e.g., making the radar direction 45° from the wall as default, as illustrated in Fig. 4(a).

Step 2 - User Centroid Estimation. The users' positions relative to the room coordinate system at this moment are consistent with the current positions calculated by the sensor. Then, we adopt a k-means clustering algorithm to look for the centroid of the users with their positions as the feature. At last, the sensor is adjusted to point to the centroid for capturing the reflected signals from the users. This adjustment process is controlled by an auxiliary rotating device, which will be described in Section IV-C.

Step 3 - Coordinate System Converting. As illustrated in Fig. 4(a), we consider the case that people might change their locations, and their centroid will also change accordingly. We describe the case in Fig. 5(a), after User 3 and User 4 move to the new locations, the system repeats Step 2 to estimate the current centroid of the users and rotates the sensor to point at it afterward. In this operation, the major challenge is that we cannot directly calculate the centroid for the current cluster since location measurements for User 3 and User 4 are relative to the sensor coordinate system (i.e., $X_s - Y_s$), whereas the locations of User 1 and User 2 are corresponding to the room (i.e., $X_r - Y_r$). To address this issue, we develop a mapping relationship between the two coordinate systems, as shown in Fig. 5(b). The mapping problem can be defined as follows:

- ***Condition:*** *Given user P's location measurement $P_s(\theta_s, d_s)$ in $X_s - Y_s$, rotating X_sY_s axes counterclockwise through an angle of φ into X_rY_r axes.*
- ***Resolve:*** *Determine the translation rule $\mathbf{T}(\varphi)$ to make the equation $\overrightarrow{OP_r} = \mathbf{T}(\varphi)\overrightarrow{OP_s}$ true.*

According to the measurement $P_s(\theta_s, d_s)$, we have $P's$

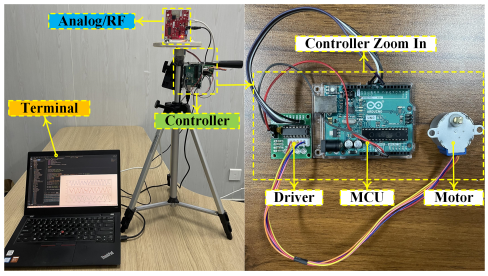


Fig. 6. The hardware setup.

coordinate $P_s(d_s \sin\theta_s, d_s \cos\theta_s)$ in the sensor coordinate system. Then, $P_r(x_r, y_r)$ in the room coordinate system can be calculated as:

$$\begin{cases} x_r = OB + \underset{=AD}{BC} = d_s \sin\theta_s \cos\varphi + d_s \cos\theta_s \sin\varphi \\ y_r = OF - \underset{=HG}{EF} = d_s \cos\theta_s \cos\varphi - d_s \sin\theta_s \sin\varphi \end{cases}. \quad (5)$$

Vector $\overrightarrow{OP_r}$ can be represented in matrix form as:

$$\underbrace{\begin{bmatrix} x_r \\ y_r \end{bmatrix}}_{\overrightarrow{OP_r}} = \underbrace{\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}}_{\mathbf{T}(\varphi)} \underbrace{\begin{bmatrix} d_s \sin\theta_s \\ d_s \cos\theta_s \end{bmatrix}}_{\overrightarrow{OP_s}}. \quad (6)$$

Using Euler's Formula, $\mathbf{T}(\varphi)$ can be further simplified as:

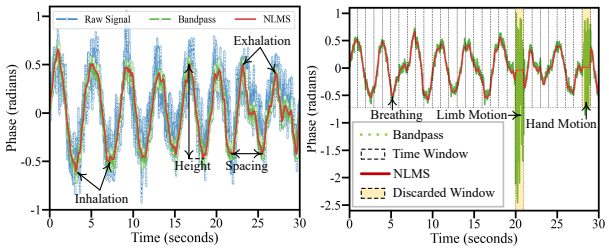
$$\mathbf{T}(\varphi) = \cos\varphi \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sin\varphi \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = exp(\varphi \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}). \quad (7)$$

Since φ is the angle between the two coordinate systems that is obtained in the previous step, we can map the locations of User 3 and User 4 to the room with $\mathbf{T}(\varphi)$.

C. Rotating Device Design

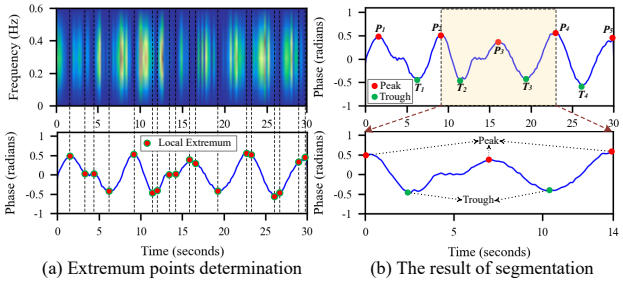
Fig. 6 shows the designed device for controlling sensor rotation. The controller primarily consists of the following parts: 1) MCU. After reading the user's location information from the radar, the MCU is programmed to calculate the centroid of the users and generate rotation instructions for the driver. 2) Driver. It is mainly composed of 4 transistors and a timer, which is in turn controlled by the MCU. The activation of the transistors provides the required voltage and current for the coils, and the timer controls its energizing timing. It controls the stepper motor in full-step driving mode, and our designed driving sequences for the motor coils are 1001, 1100, 0110, and 0011. 3) Motor. It is controlled by the clock period and rotates to the desired direction. In our implementation, we employ a unipolar stepper motor which has 5 wires one for motor supply and the other for coils.

The stepper motor provides a step angle of 1.8° and a holding torque of $3.4kg\text{-}cm$ with a $5V$ power supply, which is capable of rotating our sensor board to the desired direction. In the process of direction adjustment, our system may cause errors in user authentication due to sensor movement. This situation is likely to be a potential vulnerability that may be leveraged by attackers. In our deployment, we set the motor speed to $150rpm$ for movement stability, it just takes $0.1s$ to rotate an angle of 90°. It is almost impossible for attackers to perform malicious activities in such a short period of time.



(a) Signal filtering and smoothing (b) Body movements elimination

Fig. 7. Illustration of signal noise removal.



(a) Extremum points determination (b) The result of segmentation

Fig. 8. Respiration segment extraction.

V. Signal Processing Module

A. Respiration Signal Separation

Band pass Filtering. To identify the signal that is dominated by respiration, our insight is that the respiratory frequency band typically lies between 0.2 and $0.5Hz$, we thus adopt a Butterworth band-pass filter to cancel the irrelevant signals that are out of this band. After band-pass filtering, the signal (i.e., green-colored) shown in Fig. 7(a) shows a much higher resolution than the raw signal.

Smoothing. Since unpredictable low-frequency interference is likely to fall into the frequency range of respiration, we need further refine the robustness against impulsive interference. Specifically, we use a normalized least mean square (NLMS) adaptive filter to smooth the respiratory waveform due to its capacity of stopping the adaptive update of the filter weight in the presence of impulsive interference. From Fig. 7(a), it is observed that the signal's morphology (e.g., height and spacing) is more prominent.

Outlier removal. When a user drinks or uses the phone, body motions may cause irregular signal changes as illustrated in Fig. 7(b). Since the abrupt outbursts are aperiodic and their amplitudes are larger than respiration, it is insufficient to only rely on filters to reject them. To reduce the impact of large movements, we calculate the signal energy for a certain time window. Specifically, we slide the window over the signal, calculate the energy for each window (i.e., $\int_t^{t+1} s^2(t)dt$), and check whether it is sufficiently stronger than the signal's historical average. If the energy exceeds a specific threshold, we determine that the window is not dominated by respiration and discard it from the time-domain signal. Empirically, we use a time window of 1 second and a threshold of 5 times the energy historical average.

B. Respiration Segmentation

To facilitate feature extraction from respiration, we divide the time-domain signal into segments according to its cycles.

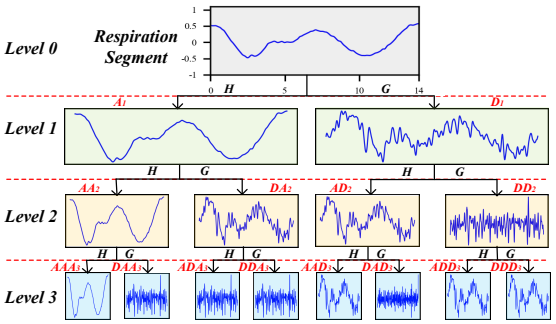


Fig. 9. Illustration of wavelet packet decomposition.

The most direct way to determine the cycles is to locate the peaks and troughs of the signal. As shown in Fig. 8(a), local extremums can be estimated by spectral analysis [12]. However, it is observed that multiple local extremums might be generated on the same peaks/troughs. To determine the exclusive points that denote the peaks/troughs in the signal, we present a distance restriction method as follows:

Extremum Classification. The peak represents exhalation that leads to a positive value in phase changes, while the trough stands for inhalation that results in a negative value. Based on this prior knowledge, we sort the extremums into maximums (i.e., positives) and minimums (i.e., negatives).

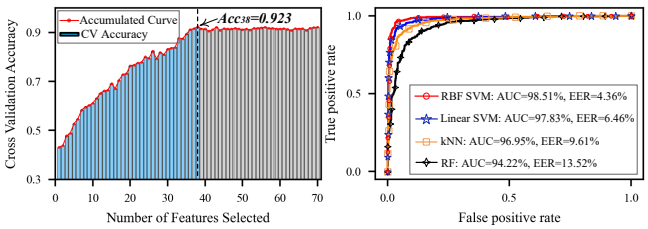
Threshold Calculation. We introduce two thresholds T_{max} and T_{min} to select the unique peaks and troughs from the two categorized groups, respectively. In particular, the thresholds are the average distances between every two adjacent values in the two groups, respectively. The average distance is calculated as: $\frac{1}{n-1}\sum_{i=1}^{n-1} t_i \times s$, where $n-1$ is the number of intervals in the group, t_i refers to the duration of the $i'th$ interval, and s denotes the sampling rate.

Peak/Trough Determination. We choose the first local maximum/minimum in the groups as a valid peak/trough, and the next valid peak/trough is selected such that the distance between the current local maximum/minimum and the previous valid peak/trough is greater than T_{max}/T_{min}. Using such restriction, we finally obtain the corresponding peaks and troughs for the signal, as shown in Fig. 8(b).

In our implementation, we slice two cycles as a respiration segment, e.g., the signal starting from P_2 to P_4 as illustrated in Fig. 8(b). The determination of cycles for a segment is further studied in Section VIII-A.

C. Biometric Features Extraction

Respiration is time-varying and non-stationary, it contains instantaneous details which are not readily obtained intuitively. To capture the representative biometrics, we employ wavelet packet decomposition (WPD) [13] to perform multi-resolution analysis in different frequency domains, facilitating us to discern the subtle differences in respiration motions between individuals. Specifically, we design a 3-level WPD with $db1$ Daubechies wavelet as illustrated in Fig. 9. The segment is decomposed into detail D and approximation A components with corresponding high-pass G and low-pass H filters at each level, respectively. The original segment is zoomed-in level by level, producing a total of 14 subspaces. At each level, every



(a) Feature selection (b) Classifier selection

Fig. 10. Feature and classifier selection.

subspace covers a part of the original frequency spectrum and thus is conducive to learning distinctive features.

For each subspace, we empirically apply 5 domain features including skewness, kurtosis, shape factor, impulse factor, and RMS. We end up with 14×5=70 features to represent the respiration segment, which are subsequently used for user template profiling and authentication model training.

VI. Authentication Module

Biometric Template Profiling. From Fig. 9, we observe that there exist duplicate components after the WPD process (e.g., DA_2 and AD_2) which may produce same features. To eliminate repetitive and less informative features, we further use the recursive feature elimination (RFE) method [14] to analyze the extracted features. Specifically, we use a linear kernel SVM classifier as the estimator for RFE. The classifier is trained by starting with all 70 features via 5-fold cross-validation. Then, RFE ranks the features by importance, discards the least important features, and re-fits the classifier. This process is recursively repeated until a specified number of features remains that make the classification reach a desired accuracy. We visualize the result in Fig. 10(a), it is observed that the curve jumps to an accumulated accuracy of 92.3% when 38 informative features are captured, then stays saturated even if choosing more features. The result reveals that the first 38 features are capable to represent respiration motions and the remaining features are not sensitive to the classification task. Based on this observation, we reduce the initial feature set to 38 features and use them to train the matching model.

User Pattern Matching. To select the optimal classifier, we compare four machine learning techniques: Random forest (RF), k nearest neighbors (kNN), linear kernel-based support vector machine (Linear SVM), and radial basis function-based support vector machine (RBF-SVM). Parameters for each classifier are tuned via 5-fold cross-validation and grid search [15] to achieve the best performance. In particular, we randomly pick one participant as the legitimate and implement a one-vs-rest strategy [16] to evaluate the performance of the classifiers. Fig. 10(b) shows the ROC curves of the four classifiers, we observe that RBF-SVM has the largest area under the curve (AUC) of 98.51% and the lowest EER of 4.36%, indicating that it is the best option for our scenario.

VII. System Setup and Data Collection

System Setup. As shown in Fig. 11, we use a COTS IWR1443BOOST mmWave radar [17] to demonstrate the

TABLE I
MMWAVE CONFIGURATION.

Bandwidth	$4GHz$	ADC Sampling Rate	$2.5M/s$
Chirp Slope	$53MHz/\mu s$	Chirp Repetition	$184\mu s$
Chirps per Frame	128	Samples per Chirp	128

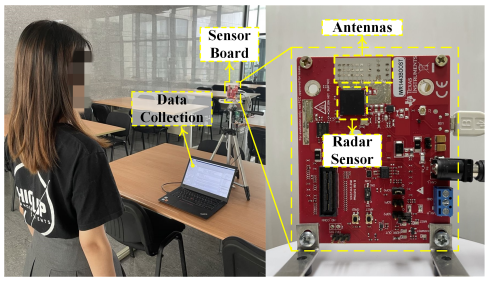


Fig. 11. Experimental setup.

feasibility of *M-Auth*. The configuration of the radar board is shown in Table I, which provides a range resolution of $3.75cm$ and a displacement resolution of $1mm$. The sensor board transmits and receives signals, then the captured signals are streamed out via UART to the laptop for further processing.

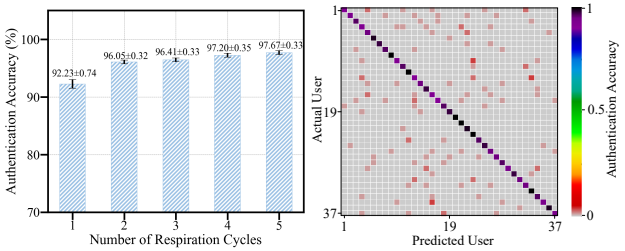
Genuine Data Collection. We recruit 37 healthy participants (17 females and 20 males) whose ages range from 19 to 35. In a typical workplace setting as illustrated in Fig. 11, each participant is asked to sit/stand in front of the radar at a distance of $2m$ and breathe naturally without any restrictions (e.g., they can move limbs and operate smartphones). The default settings are used unless stated otherwise. To maintain data persistence, data collection is done in multiple rounds for a period of two months. We collect 400 respiration segments from each participant and obtain 14,800 samples in total.

Attack Data Collection. *1) Blind Attack:* We choose 7 participants as victims and the remaining 30 participants as attackers. For each victim, every attacker arbitrarily performs 20 segments, producing 4200 samples in total. *2) Impersonation Attack:* We invite 7 participants as victims and 10 participants as attackers to mimic the victims' breathing. The attackers are asked to observe the victim's respiratory rhythm and breathing depth in a close-up view. For every victim, we obtain 50 segments from each attacker and 3500 samples in total. *3) Replay Attack:* We invite 7 participants as victims and employ an additional mmWave radar to record the victim's respiration signals. It is assumed that the end device is secure, the attacker does not know the specifications such as the length of respiration segment and the configuration of FMCW chirps. For each victim, we use the default chirp configuration to capture respiration for 10 minutes and slice the signal into segments by 5 seconds. In total, we collect 840 samples.

VIII. PERFORMANCE EVALUATION

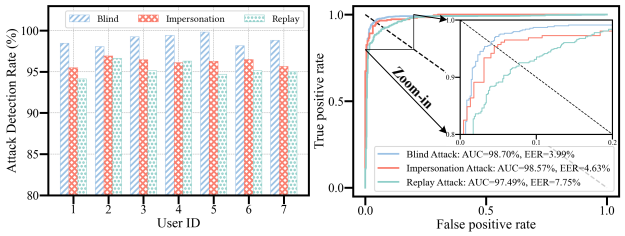
A. Overall System Performance

We first determine the appropriate respiration cycles for the segment. Fig. 12(a) shows the authentication accuracy with different respiration cycles. We observe that when 2 cycles are chosen, the average accuracy leaps to 96.05% and then remains roughly stable. The standard deviation (STD) decreases from 0.74% to 0.32% and the subsequent changes



(a) Performance with different cycles (b) Confusion matrix

Fig. 12. Performance of user authentication.



(a) Attack Detection (b) ROC curves under attacks

Fig. 13. System performance under three attacks.

are not significant. A low STD shows that the performance is more reliable, i.e., the results are clustered closely around 96.05%. Based on this observation, we slice 2 cycles for the segment and the corresponding average accuracy demonstrates that *M-Auth* is effective in verifying legitimate users.

Next, we evaluate the performance of specific user verification. Fig. 12(b) shows the confusion matrix for the 37 participants. It presents the corresponding authentication accuracy along the diagonal regions. Among all the participants, the lowest and the highest authentication accuracy are 92.25% and 100%, respectively. A darker area denotes higher accuracy, validating the reliability in identifying individuals.

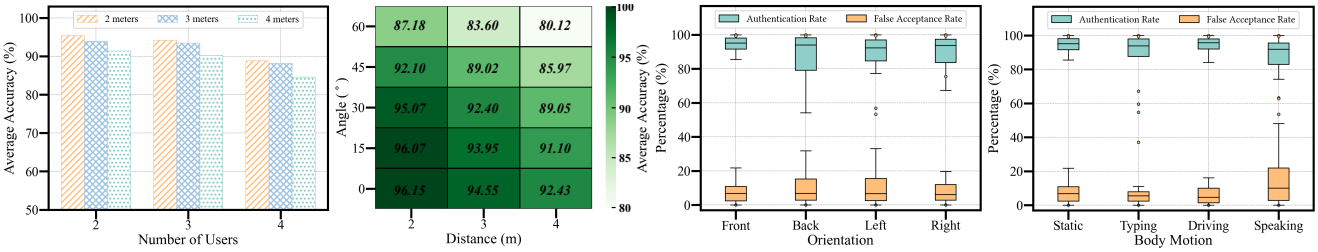
B. Performance of Resisting Attacks

We evaluate the resilience of *M-Auth* for the attacks discussed in Section II-C. As shown in Fig. 13(a), the detection rates of the three attacks are over 98%, 95%, and 94%, respectively, with mean values of 98.86%, 96.29%, and 95.34%, respectively. It is expected to have a high detection rate under blind attack since respiration motions are rarely the same between individuals as mentioned in Section II-A. A slight decrease in detection of impersonation attack implies that imitating the victim's respiration helps the attack; however, precise replication is challenging to accomplish. Besides, as attackers lack detailed specifics of *M-Auth*, such as chirp configuration and signal segmentation, our system maintains its resilience to replay attack as verified by the results.

The ROC curve in Fig. 13(b) provides a more intuitive understanding of the system performance. Specifically, the AUCs under the three attacks are 98.70%, 98.57%, and 97.49%, respectively; and the EERs are 3.99%, 4.63%, and 7.75%, respectively. Higher AUC and lower EER represent that our system is effective to distinguish legitimate users from attacks.

C. Robustness Analysis

Impact of Multiple Users Under Variant Distances. We evaluate our system with up to four users under a distance



(a) Multiple users under different distances (b) AoAs under different distances (c) Different user orientations (d) Different body motions

Fig. 14. Robustness analysis for different scenarios.

varied from $2m$ to $4m$. We randomly ask 2, 3, and 4 of our participants to stand shoulder to shoulder at a distance of $2m$, $3m$, and $4m$, respectively. Fig. 14(a) shows the average authentication accuracy. It is observed that the accuracy values are all over 90% for the groups of 2-user and 3-user within $4m$. The accuracy values for the 4-user within $3m$ almost approach 90%. The results verify that *M-Auth* is capable of identifying multiple users simultaneously within a reasonable range. Note also that the increase of either users or the distance might decrease the accuracy, especially for the group of 4-user at $4m$, the accuracy decreases to 85%. It is expected due to the mmWave property of fast attenuation and could be further improved by increasing respiration cycles in the segment.

Impact of AoA Under Variant Distances. To examine the effective sensing range, we evaluate the performance under changeable AoAs and distances. Specifically, the recruiter is required to sit/stand at $2m$, $3m$, and $4m$ from the system, respectively, and at angles ranging from $0°$ to $60°$ with respect to the radar orientation. The results are visualized in Fig. 14(b). We observe that the accuracy exceeds 92% when the angle is less than $30°$ and the distance is within $3m$. Like the previous experiment, the accuracy decays with the increase of distance. Moreover, the accuracy is above 92% when the user is on a straight line with the probe and reduces to less than 88% at $60°$. The results are expected since the estimation of phase change decays with the increase of AoA. This experiment motivates us to design the rotating device in Section IV-C, which can dynamically adjust the device orientation according to the user position.

Impact of User Orientation. We study the performance when users do not directly face the device. Participants are asked to perform four different orientations: facing the device, having their back to the device, and facing left or right to the device. We provide the results in Fig. 14(c). When facing the device, we observe that the average authentication and false acceptance rates are the best (96.07% and 6.23%, respectively). Across all the orientations, they slightly fluctuate by 1%-3%, which shows the robustness of our system to verify the user in different orientations. This is because when one breathes, the chest expands in all directions, and *M-Auth* can capture the side expansions.

Impact of Body Motion. We further investigate the performance under daily activities without requiring users to stop their ongoing work. Participants are invited to perform four different activities: static (as a control group), typing, imitating driving and speaking. From Fig. 14(d), it is observed that the average authentication rate and false acceptance rate are close to those of the control group when typing or driving. The results are consistent with our methodology, where we introduce a signal energy comparison scheme to remove the outliers caused by limb or hand movements. In the case of speaking, the results drop by about 6% compared with the control group. This is due to the inherent nature of speaking, i.e., phonation relies on exhalation; it is not possible to phonate during inhalation [18]. The limitation could be alleviated by intermittent pauses during speaking.

IX. Related Work

Continuous Authentication. Traditional physiology-based authentications, such as fingerprint [19], iris [20], and face [21], only provide a one-time authentication at the start of a login session and are vulnerable to artifacts. To prevent security concerns, behavior-based continuous authentications including gait patterns [1], vocal vibrations [22], and keystroke dynamics [2] are proposed. These works, however, require users to continuously and actively interact with the authentication system, which is obtrusive and not user-friendly in practical use.

Vital Sign-based Authentication. To reduce the limitations above, vital signs are exploited to seek novel passive authentications. Specifically, brain responses produced by visual stimuli [23], ECG [24], and PPG [25] are studied for continuous authentication. However, these studies have the limitation of requiring users to wear a body-attached gadget, which is inconvenient and restricts their application scenarios. To perform a non-contact manner, wireless signals have been targeted to detect vital signs for continuous authentication. For example, Lin *et al.* use a CW Doppler radar to sense the unique cardiac motion for verifying users [3]. These methods, however, either implement with a dedicated device or require close-range sensing, largely affecting their applications.

The most related work to ours is *BreathID* [4]. The authors use the deployed WiFi signal to sense the unique respiratory motions for authentication. The remarkable factors that distinguish this work from ours lie in: 1) Their sensing modality limits them to verify one person at a time, while we can test multiple users concurrently. 2) They assume users are stationary, ignoring limb and hand motions. By eliminating motion-corrupted segments, we make the system more practical. 3) They choose more than 500 features to describe the respiration signal. We eliminate redundant and misleading features and elaborately identify 38 representative

ones, which benefits authentication accuracy and user template update. With these unique factors, we provide a different breed of authentication from *BreathID*.

X. Discussion

We analyze the potential limitations of our system and provide suggestions for how to improve it in the future.

Exercising and Health. In our work, we use respiration data acquired from healthy people under normal physical conditions to construct the matching model. People who have just finished exercising or have breathing problems, such as asthma, pneumonia, and anxiety, may have significant fluctuations in respiration motions, affecting the authentication accuracy. One potential method for improving the resilience of our system is to determine how sensitive *M-Auth* is to such contextual changes. For example, we may study how quickly a user's respiration rate drops after exercise and utilize the respiration recovery rate as one of the matching features.

Quasi-static State. To avoid large body movements from producing system errors, our system requires the users to be quasi-static. This constraint is a common stumbling block for wireless sensing; because phase shifts caused by full-body motions often drown those caused by respiration, preventing the monitoring of minute skin fluctuations. Isolating the submerged signal is not trivial due to its low signal-to-noise ratio (SNR). To apply *M-Auth* to full-body movement contexts, one feasible solution is to employ intermittent authentication, in which users are requested to pause their current activities for a short time for authentication.

XI. Conclusion

This paper presents a continuous multi-user authentication system by sensing respiratory motion using a single COTS mmWave radar. We design a rotating device to help the radar obtain high-quality reflected signals from users. To effectively identify legitimate users and thwart spoofing attacks, we experimentally determine the appropriate data segment, elaborately select the representative features, and build a fine-tuned classifier. Extensive experiments demonstrate that our system is resilient to spoofing attacks and effective in authenticating legitimate users in various application scenarios.

Acknowledgment

This work is supported by NSFC (Grant No. 62002278).

References

[1] Y. Z. Zeng, P. H. Pathak, and P. Mohapatra, "Wiwho: Wifi-based person identification in smart spaces," in *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2016, pp. 1–12.

[2] J. Kim and P. Kang, "Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features," *Pattern Recognition*, vol. 108, 2020.

[3] F. Lin, C. Song, Y. Zhuang, W. Xu, C. Li, and K. Ren, "Cardiac scan: A non-contact and continuous heart-based user authentication system," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking (MOBICOM)*, 2017, pp. 315–328.

[4] J. Liu, Y. Chen, Y. Dong, Y. Wang, T. Zhao, and Y. D. Yao, "Continuous user verification via respiratory biometrics," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2020, pp. 1–10.

[5] H. Kong, L. Lu, J. Yu, Y. Chen, X. Xu, F. Tang, and Y.-C. Chen, "Multiauth: Enable multi-user authentication with single commodity wifi device," in *Proceedings of the International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MOBIHOC)*, 2021, pp. 31–40.

[6] M. Muaaz and R. Mayrhofer, "Smartphone-based gait recognition: From authentication to imitation," *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3209–3221, 2017.

[7] S. Yue, H. He, H. Wang, H. Rahul, and D. Katabi, "Extracting multi-person respiration from entangled rf signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 2, no. 2, pp. 1–22, 2018.

[8] F. Adib, H. Z. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 2015, p. 837–846.

[9] S. M. Islam, N. Motoyama, S. Pacheco, and V. M. Lubecke, "Non-contact vital signs monitoring for multiple subjects using a millimeter wave fmcw automotive radar," in *2020 IEEE/MTT-S International Microwave Symposium (IMS)*. IEEE, 2020, pp. 783–786.

[10] C. Iovescu and S. Rao, "The fundamentals of millimeter wave sensors," *Texas Instruments*, pp. 1–8, 2017.

[11] S. Rao, "Introduction to mmwave radar sensing: Fmcw radars," *Texas Instruments*, pp. 1–70, 2020.

[12] F. Scholkmann, J. Boss, and M. Wolf, "An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals," *Algorithms*, vol. 5, no. 4, pp. 588–603, 2012.

[13] R. X. Gao and R. Q. Yan, *Wavelet Packet Transform*. Boston, MA: Springer US, 2011, pp. 69–81.

[14] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sensors and Actuators B: Chemical*, vol. 212, pp. 353–363, 2015.

[15] I. Syarif, A. Prugel-Bennett, and G. Wills, "Svm parameter optimization using grid search and genetic algorithm to improve classification performance," *Telkomnika*, vol. 14, no. 4, p. 1502, 2016.

[16] Y. Xue and M. Hauskrecht, "Active learning of multi-class classification models from ordered class sets," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 5589–5596.

[17] T. Instruments, "Iwr1443 evaluation module (iwr1443boost) mmwave sensing solution user's guide," https://www.ti.com/tool/IWR1443BOOST, 2020, accessed May 19, 2020.

[18] Y. Wang, W. D. Cai, T. Gu, W. Shao, Y. N. Li, and Y. Yu, "Secure your voice: An oral airflow-based continuous liveness detection for voice assistants," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 3, no. 4, 2019.

[19] O. Ogbanufe and D. J. Kim, "Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment," *Decision Support Systems*, vol. 106, pp. 1–14, 2018.

[20] H. Shahriar, H. Haddad, and M. Islam, "An iris-based authentication framework to prevent presentation attacks," in *2017 IEEE 41st annual computer software and applications conference (COMPSAC)*, vol. 2. IEEE, 2017, pp. 504–509.

[21] R. Raghavendra, K. B. Raja, and C. Busch, "Presentation attack detection for face recognition using light field camera," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1060–1075, 2015.

[22] H. N. Li, C. H. Xu, A. S. Rathore, Z. X. Li, H. B. Zhang, C. Song, K. Wang, L. Su, F. Lin, K. Ren, and W. Y. Xu, "Vocalprint: exploring a resilient and secure voice authentication via mmwave biometric interrogation," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems (SENSYS)*, 2020, pp. 312–325.

[23] F. Lin, K. W. Cho, C. Song, W. Y. Xu, and Z. P. Jin, "Brain password: A secure and truly cancelable brain biometrics for smart headwear," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (MOBISYS)*, 2018, pp. 296–309.

[24] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, "Ecg authentication for mobile devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 591–600, 2015.

[25] Y. M. Chen, J. C. Sun, X. C. Jin, T. Li, R. Zhang, and Y. C. Zhang, "Your face your heart: Secure mobile face authentication with photoplethysmograms," in *in Proceedings of IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2017, pp. 1–9.