BioCase: Privacy Protection via Acoustic Sensing of Finger Touches on Smartphone Case Mini-Structures

Yilin Yang
Rutgers University
yy450@scarletmail.rutgers.edu

Xin Li
Rutgers University
xl658@scarletmail.rutgers.edu

Zhengkun Ye
Temple University
zhengkun.ye@temple.edu

Yan Wang
Temple University
y.wang@temple.edu

Yingying Chen
Rutgers University
yingche@scarletmail.rutgers.edu

ABSTRACT

Finger biometrics are widely used by smartphones as a secure and user-friendly credential for privacy protection. However, this information is difficult to measure without high-resolution images, leaving most works to treat this as an image-domain problem. We demonstrate that low-effort alternatives on smartphones are possible through the use of sound propagation in ubiquitous smartphone cases. Inexpensive and widely adopted, smartphone cases are always in contact with fingers, making them ideal for collecting finger biometrics. We thus design BioCase, an acoustic sensing system that leverages smartphone cases equipped with mini-structures to capture unique biometric-hybrid signatures (i.e., reflections influenced by the user's fingertip physiology and behavior) for smartphone privacy protection. The system generates inaudible structure-borne sound and measure the propagation through the smartphone case, mini-structures, and user finger. The design of the mini-structure controls the behavior of structure-borne sound such that unique responses are produced when different users and fingers touch the smartphone case. This enables low-cost, low-effort privacy protection, merely touching the smartphone case can authenticate users. Comprehensive experiments with 46 users over 10 weeks demonstrate BioCase can differentiate users with over 94% accuracy at a 5% false positive rate.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**.

KEYWORDS

Privacy Protection, Acoustic Sensing, Smartphone Cases

ACM Reference Format:
Yilin Yang, Xin Li, Zhengkun Ye, Yan Wang, and Yingying Chen. 2023. BioCase: Privacy Protection via Acoustic Sensing of Finger Touches on Smartphone Case Mini-Structures. In *The 21st Annual International Conference on Mobile Systems, Applications and Services (MobiSys '23), June 18–22, 2023, Helsinki, Finland*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3581791.3596841

MobiSys '23, June 18–22, 2023, Helsinki, Finland

ACM ISBN 979-8-4007-0110-8/23/06…$15.00
https://doi.org/10.1145/3581791.3596841

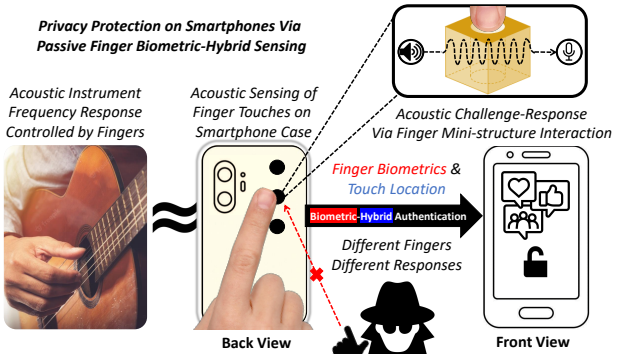


Figure 1: BioCase prevents privacy leaks by verifying acoustic responses produced by user fingers when touching smartphone cases before displaying alerts.

1 INTRODUCTION

Smartphones have become invaluable tools for daily work and recreation. However, they are also high-priority targets for attackers due to the sensitive data they carry. Information such as text messages, work contacts, social media, photos, and health app records are all concentrated on a single device that is prone to theft or snooping. User authentication is often used to secure sensitive content. Knowledge-based methods (e.g., password, PIN) are low-cost but require cognitive and physical effort (i.e., memorization, typing long strings). Biometric-based authentication lowers this burden by measuring user physiology, such as face or voices. However, this creates dependency on specialized sensors. This motivates two of our goals: (1) low-effort biometric-based privacy protection usable anytime and anywhere while (2) lifting the authentication burden from specialized senors.

We observed that smartphone owners often use protective cases for their devices [15, 18]. These cases are highly diverse in shape, material, elasticity, and other attributes that can modify the structural characteristics of the smartphone. They may also contain holes or hollow chambers for shock absorption, creating *mini-structures* in the greater smartphone body. This is particularly interesting when measuring *structure-borne sound*. When griping the phone, the user's fingers naturally touch the back surface, influencing the propagation behavior of structure-borne sound due to the combined medium of the fingers, smartphone, and case. Properties such as

finger and mini-structure shape strongly affect acoustic signals in the form of propagation speed, acoustic diffraction, and signal attenuation. Well designed mini-structures can even alter oscillation of acoustic frequencies, resulting in amplified frequency responses. Such responses manifest as resonant frequency and are commonly observed in musical instruments. In this work, we propose to leverage this phenomenon to recognize smartphone users by their finger touches.

The majority of finger sensing works favor image domain approaches, which require high quality scans and cameras [10, 22, 36]. Unfortunately, finger scans can be stolen or presented without consent [40, 44]. Scanning only works at a single pre-determined sensor location, which can be obstructed by dirt or sweat. We envision an alternative where users are free to touch anywhere on the smartphone in order to authenticate themselves and control their private data (e.g., hide/display message previews, incoming phone calls).

A few prior works have explored using acoustic sensing for recognizing users, including face scans [45], breathing patterns [4], and lip movements [24]. These works require high-effort posing, gestures, or speech utterances. Recently, several pioneering works have tried to capture palm biometrics on smartphones [12, 42]. However, palm biometrics are increasingly under threat by sophisticated attacks (e.g., 3D printed hands [25], gloves [9]). Unlike previous works, we use acoustic sensing to recognize users based on their fingertip physiology and behavior traits, referred to as *biometric-hybrid signatures*, which aim to provide more security and flexibility for privacy protection on smartphones. Acoustic sensing is capable of distinguishing coarse-grained finger touch locations [34], but to our knowledge, no works are sufficiently fine-grained to identify users with finger touches. We find that mini-structures of smartphone cases can enhance the sensitivity of structural-acoustic sensing, differentiating users not only by finger, but also by finger quantities and touch locations. We exploit this in our biometric-hybrid signature to verify that knowledge-based keys (e.g., touch location) can only be accepted if inputted by a valid user (e.g., authorized finger), preventing knowledge-based attacks with no further effort from the user [1, 23]. Similarly, biometric attacks cannot succeed without knowledge of a secret key (e.g., touch location, finger to use, number of fingers). Finger touches on smartphone cases happen at the back of the phone, making them difficult to observe through popular shoulder surfing attacks.

Motivated by these observations, we propose BioCase, a novel privacy protection system that leverages mini-structures of smartphone cases to produce structure-borne acoustic responses unique for different users and fingers. Intuitively, the combined structure of the user finger, smartphone, smartphone case, and internal mini-structures creates a uniquely shaped propagation medium difficult to replicate. Much like acoustic instruments, different shaped structures can produce different acoustic frequency responses controllable by finger interactions as depicted in Figure 1. BioCase is accessible in practice because smartphone cases are low-cost and common (reports for 2017 estimate 79% of users have cases [19] while only 55% have fingerprint scanners [37]). We embed smartphone cases with multiple mini-structures that each produce distinct resonant frequencies, providing *frequency diversity*. Since mini-structures are located at different positions along the case, our responses also provide *spatial* diversity.

We stimulate the smartphone body and elicit such responses via a high-frequency acoustic challenge signal using built-in speakers. By pressing one or more fingers against mini-structures, users can alter the resonance of the challenge and produce a unique acoustic response. The force exerted by the finger changes the mini-structure shape such that the frequency oscillation will be changed, thereby producing different resonant frequencies, akin to pressing keys on a musical instrument. The changed frequency response is measurable by built-in microphones and is specific to both the finger that triggered the compression and mini-structure being compressed. Attackers cannot impersonate the response without a matching finger structure and knowledge of which mini-structure to press, making our biometric-hybrid signature more robust than biometrics or knowledge alone. Eavesdropped responses cannot be deployed due to loss of structural information in in-air recordings. To defeat replay attacks using engineered signals, we randomize qualities of our acoustic challenge signal at each transmission. Attackers cannot predict the qualities of the randomized challenge and fabricate the necessary response in real time.

This prevents information such as text messages, contact lists, social media, photos, and health monitoring records from being easily leaked on screen. Lock screens frequently display notifications that expose data without any authentication check. Disabling the display of such alerts prevents privacy leaks, but also denies the user access to their own information. BioCase can greatly enhance the convenience and security of privacy protection on smartphones by concealing sensitive onscreen information until unlocked by a simple finger touch. Incoming text messages, caller IDs, and emails can be selectively hidden until touches from an authorized finger are confirmed. This allows for seamless, real-time protection of private information with unpredictable display behaviors. We envision BioCase can pioneer a new pillar of authentication techniques alongside existing methods like Face ID [13] or Touch ID [6]. BioCase can also work together with existing methods through multi-factor authentication as a low-effort enhancement to smartphone security. We studied the feasibility of achieving such privacy protection using finger touches on smartphone cases and made the following contributions:

- We propose a novel privacy protection system, BioCase, that authenticates users via finger touches on smartphone cases instead of specialized sensors. Finger touches embed user-specific information into structure-borne sound propagating in the smartphone case, allowing us to distinguish users and selectively display or hide their private data.
- We design a biometric-hybrid signature leveraging the user's natural interactions with the smartphone case (i.e., touches when holding the case). This signature is more convenient than knowledge credentials (e.g., PINs) alone since finger touches require less effort than typing/swiping. It is also more robust than biometrics alone, as attackers cannot easily reproduce the user's finger touch behavior. Moreover, it is more secure than either of them because our signature is unique for each user in knowledge credentials and finger physiology.

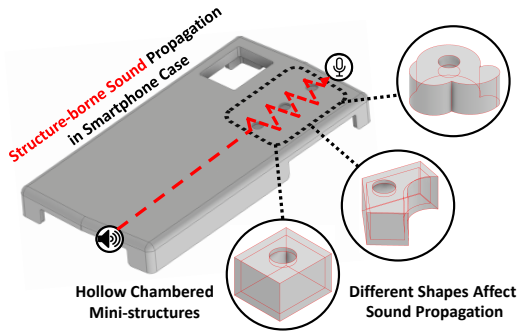


Figure 2: Sound propagation inside a smartphone case.

- We study mini-structures and how their different specifications can produce different resonant frequencies in smartphone cases. We apply our findings towards developing an acoustic response embedded with frequency-diversity and user finger touch information that is reliable for authentication and difficult to attack.
- We conduct extensive experiments involving 46 participants, multiple smartphone devices, and smartphone cases. Our experiments consider real-world factors including noise, movement, and user finger interactions. Results show that users can be recognized with over 94% accuracy.

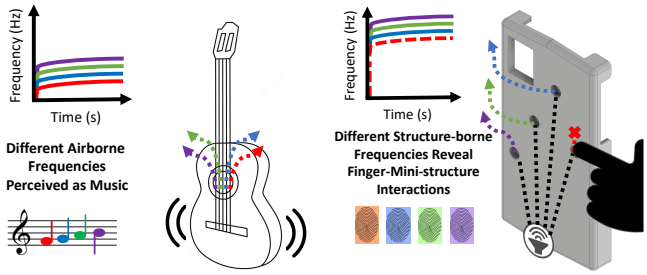
2 ACOUSTICS OF MINI-STRUCTURES

We introduce the behavior of acoustic signal propagation in physical structures relative to external forces caused by user hands and internal forces controlled by mini-structures.

2.1 Spatial Diversity in Mini-structures

We provide privacy protection by verifying users through their finger touches as they operate their smartphones. In particular, we utilize acoustic information to conduct tactile sensing. Unlike common acoustic-based technologies, which leverage airborne propagation to distinguish user voices or gestures, we utilize structure-borne sound propagation and its behavior in differently shaped objects. Most smartphones have well-defined dimensions (i.e., rectangular prism) and relatively uniform density, producing consistent structural acoustic responses. When touching the smartphone, the user's hands and fingers transform the shape of the structure and the resulting response. Moreover, different users produce unique responses due to their distinctive finger shapes, tissue structure, and bone density. The smartphone body and user finger therefore contribute to a *joint structure* that is specific to the user-smartphone combination. We find that, rather than impose burden on the user (i.e., extensive hand motions or gestures), we can generate the unique acoustic responses by controlling the joint structure.

Toward this end, we propose to *expand* the joint structure through the use of smartphone cases as shown in Figure 2. We embed internal hollow mini-structures in the case to intentionally change the acoustic response based on structural specifications. Mini-structures are linked to the case surface through connecting tubes, creating holes that the user can easily identify and touch. We illustrate this concept in Figure 3. Each mini-structure is shaped differently in order to produce different responses based on resonant frequency. We liken this concept to frequency responses in musical instruments in Figure 3(a). Stimuli such as string vibrations produce different



(a) Conventional usage of airborne sound in acoustic structures.

(b) BioCase usage of structure-borne sound in acoustic mini-structures.

Figure 3: Mini-structures can produce acoustic responses sensitive enough to capture finger biometric-hybrid signatures.

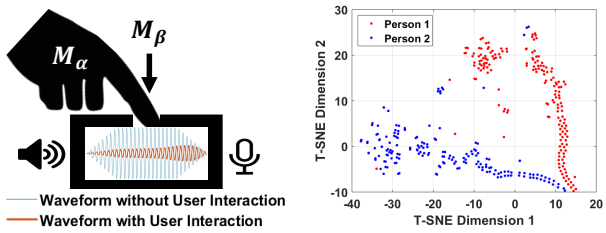
frequencies based on characteristics of the string and instrument structure. BioCase employs this concept on a smaller scale using smartphone case mini-structures in Figure 3(b). The stimulus in this scenario is a high-frequency inaudible signal. Through structural propagation, the signal will permeate into the mini-structure and resonate if the mini-structure size and pressure are sufficient to amplify oscillation, yielding two observations: (1) The small size of the mini-structure ensures that only the inaudible-frequencies will resonate. (2) Different mini-structures specifications will resonate with different frequencies, producing distinct acoustic responses. Finger touches that cover the mini-structure tubes will alter specific frequency bands. Similar to fingerprints, the altered frequencies are also user-specific due to the unique physiological traits of the finger. By observing how frequencies are altered, and knowing the location and resonant frequencies of each mini-structure, Biocase can derive biometric-hybrid signatures. Our mini-structures have minimal location constraints so long as they are in the structure-borne sound propagation path, meaning they can be placed almost anywhere on the back surface of the smartphone case. They can also be placed in large quantities, enhancing both security (i.e., increasing sensing resolution) and convenience (i.e., allowing more touch locations to authenticate the user).

2.2 Frequency Diversity in Mini-structures

Structure-borne sound behavior can be predicted knowing speed of propagation c and physical properties of the propagation medium. We model behavior within mini-structures as a form of Helmholtz resonance H:

$$H = \frac{c}{2\pi}\sqrt{\frac{A}{VL}}. \quad (1)$$

where V is volume of the internal space. A hole is bored from the mini-structure to the exterior surface of the smartphone case, the area and length of which is denoted as A and L. Similar to hollow chambers of musical instruments, the mini-structure amplifies the oscillations of different frequencies through constructive interference, creating stronger vibrations or 'echoes'. These resonant frequencies can be intentionally enhanced or diminished through the design of the mini-structures as well as user interaction. For example, sealing the hole to the mini-structure with the user finger can decrease the volume V and increase pressure in the internal chamber, absorbing sound and altering the resonance akin to soundproofed rooms. By transmitting a frequency matching the



(a) Mini-structure interaction model (not to scale). No two users will have identical responses due to their physiology.

(b) T-SNE plot of acoustic features for two users pressing the same mini-structure. Different fingers are clearly separable.

Figure 4: Modeling of user interaction with smartphone mini-structures. The unique shape and mass of the finger will alter mini-structure resonance frequencies to yield user-specific acoustic features.

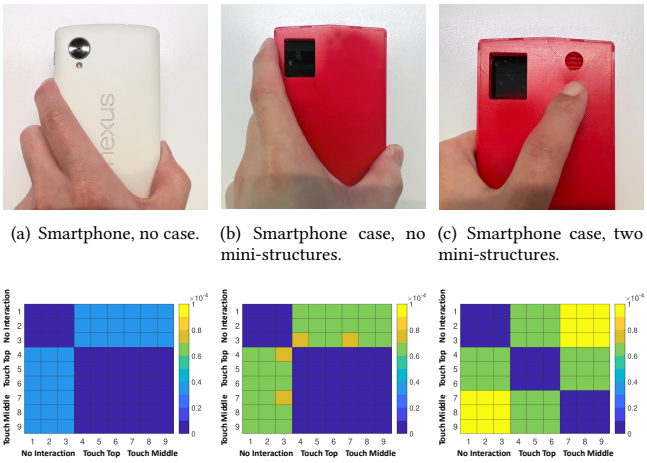
resonance of the mini-structure, it is possible to determine whether the user touched the mini-structure based on the received signal strength. For multiple mini-structures, the user response $R(\omega)$ to a transmitted signal $S(\omega)$ can be modeled as:

$$R(\omega) = \sum_{i=1}^{n} H_i M_\beta(i) S(\omega) M_\alpha, \tag{2}$$

where M_α and M_β are the forces exerted by the external mass (e.g., user hand, table, etc.) and a user finger, respectively. Note that M_β is a nonempty tuple rather than a constant like M_α as the user may use multiple fingers of differing mass to interact with the mini-structures. The response signal is therefore a unique convolution determined by the physical characteristics of the external mass, mini-structure, and user finger. Note that sealing all mini-structures simultaneously causes H_i to approach 0, allowing Eq. 2 to be approximated as:

$$R(\omega) = S(\omega) M_\alpha, \tag{3}$$

which is functionally equivalent to having no mini-structures at all and diminishes the frequency diversity. Note that different users have different M_α and M_β values as no two fingers are completely identical. When exerting force against the mini-structure, these differences embed user-specific acoustic features in the resonant frequency response, even when interacting with the same mini-structure. This model of user interaction with BioCase is illustrated in Figure 4. We elaborate on specific features in Section 5.1. A total of n mini-structures may exist, distinguished by numerical designation i. The user may choose to cover zero or more mini-structure holes with their finger, creating distinct responses for each possibility from the same user. This ensures potential attackers cannot easily forge our biometric-hybrid signature without similar biometric credentials (i.e., external mass M_α and finger M_β shapes) and knowledge credentials (i.e., finger interaction $M_{(\beta,i)}$ with mini-structure H_i). The frequency diversity of the acoustic responses can be most effectively leveraged by designing the mini-structure resonant frequencies and transmitted acoustic signal to operate within the same bandwidth.



(a) Smartphone, no case.

(b) Smartphone case, no mini-structures.

(c) Smartphone case, two mini-structures.

(d) Response to different touch interactions, no case.

(e) Response to different touch locations, case without mini-structures.

(f) Response to different touch locations, case with two mini-structures.

Figure 5: Euclidean distance of amplitudes from different touch locations. Smartphones with mini-structures produce stronger differences than without.

2.3 Feasibility Study

A preliminary experiment was performed to understand the sensitivity of structure-borne sounds to different finger touch interactions. Three scenarios were examined to demonstrate how changes to the internal structure of the propagation medium can influence structure-borne sound propagation. Figure 5 shows the experimental setup for each scenario; using a smartphone with no case, a smartphone equipped with a case and no mini-structures, and a smartphone equipped with a case and two mini-structures. For each scenario, we attempt to distinguish between no finger interaction, finger touches near the top of the device, and finger touches near the middle of the device. Three samples of each interaction are obtained, denoted as 1-3, 4-6, and 7-9 in Figure 5(d)-5(f). Note that for the scenario shown in Figure 5(c), we position mini-structures for the user to touch. The smartphone cases for Figure 5(b) and Figure 5(c) are otherwise identical. The scenario without mini-structures approximates the same touch location by using the camera lens position as a reference point. An inaudible chirp signal sweeping from 18 kHz to 22 kHz was emitted from the bottom device speakers to induce vibration, which is then recorded by a single microphone near the top of the device.

We extract the average amplitude for each recording. The Euclidean distance was then computed for each combination of samples to quantify similarities and differences. The results can be seen in Figure 5(d)-5(f). Samples native to the same scenario (e.g., 1 & 2) displayed naturally low amplitude difference whereas samples of different scenarios (e.g., 1 & 9) had increased differences. Note across the diagonal that the distance is zero as these are comparisons between a sample with itself. We observe in Figure 5(d) that the base smartphone may coarsely distinguish between the touches and no touches, but cannot adequately differentiate between different touch locations. Equipping a smartphone case, however, enhances

sensitivity by expanding the propagation medium. With the inclusion of mini-structures, this sensitivity is responsive enough to recognize the different touch location. This suggests that, even with minimal sensors and signal processing, structural-borne sound propagation is capable of communicating critical information about the user and their interaction with the smartphone device.

3 ATTACK MODEL

Attackers may wish to gain unauthorized access to personal data (e.g., email, social media, shopping accounts, health monitoring apps) on smartphones. We consider attack scenarios under the following assumptions: (1) The attacker can be in close proximity to the legitimate user such that they can monitor the user or gain physical access to the smartphone unnoticed. (2) The attacker has full knowledge of our acoustic challenge-response system and how it measures the biometric-hybrid signature. (3) The attacker does not have power over operating systems to skip authentication checks. Accordingly, we define the most feasible attacks below:

Impersonation Attack. The attacker attempts to impersonate the legitimate user's acoustic response by mimicking their hold style and finger touches on the mini-structures. We consider attackers that may either be informed or uninformed. For the uninformed case, the attacker has no information of the authentication process and can only impersonate the legitimate user through educated guessing. For the informed case, the attacker has knowledge of the legitimate user's biometric-hybrid signature via passive observations.

Replay Attack. The attacker may bypass the challenge-response by presenting a pre-obtained signal instead of their own live-generated physical response. The signal can be stolen from the legitimate user by eavesdropping instances of authentication attempts. This can be done by placing external microphones near the target smartphone or by installing a malicious app that covertly controls the internal microphones of the user's smartphone. In such scenarios, the attacker can obtain a signal embedded with valid credentials known to authenticate the user. Alternatively, an original signal can be synthesized that has similar properties to the legitimate user response. Afterwards, the attacker gains access to the smartphone device and replays the signal in guise of the legitimate user using a speaker.

Puppet Attack. The attacker may bypass the challenge-response by moving the legitimate user's hands and fingers into the positions necessary for BioCase to measure and authenticate. This attack is possible during situations where the user is unresponsive (e.g., asleep, intoxicated). The attacker must know the correct touch locations to press the user's fingers against; the knowledge of this configuration can be obtained through passive observation of the user.

4 SYSTEM DESIGN

4.1 Privacy Protection Framework

Our system must distinguish between different people to ensure private content is disclosed to authorized users only. We consider the example scenario where data may be leaked when an incoming message is displayed on the smartphone screen without regard for who can see it. To prevent this breach of privacy, we selectively hide

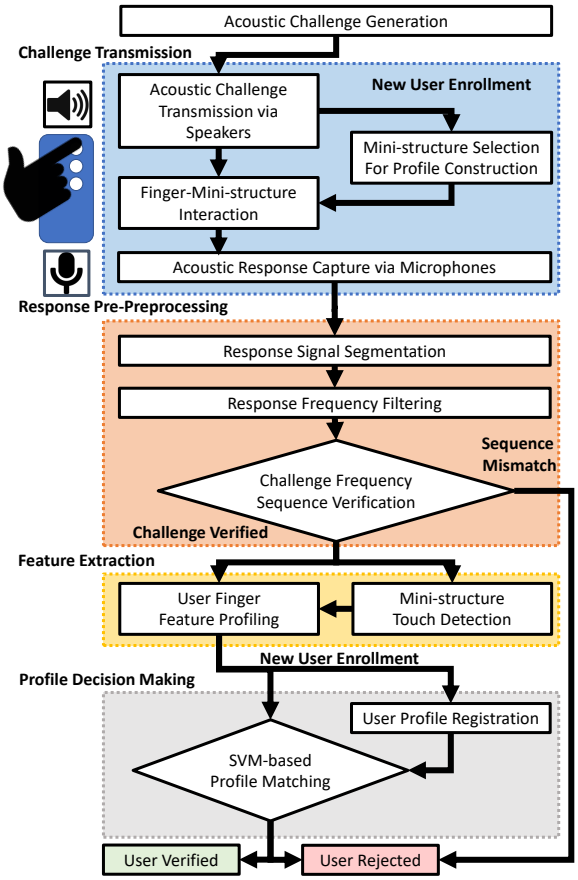


Figure 6: System overview of BioCase.

or display this content depending on the user identity touching the smartphone. BioCase provides smartphone user authentication by measuring user response to an identity challenge in the form of an acoustic signal transmission. Finger physiological and behavioral information are captured in our biometric-hybrid signature through the transformation of the acoustic challenge signal as it propagates through mini-structures. The combination of the smartphone, user finger, and smartphone case mini-structures creates a uniquely shaped medium that produces distinct features for different users. We thus propose the privacy protection system shown in Figure 6, consisting of four major components.

Challenge Transmission. A potential privacy leakage incident such as the incoming message will trigger the built-in speaker of the smartphone to transmit our acoustic challenge. A pool of inaudible frequency chirps is used to dynamically generate a randomized challenge at each transmission instance. The specific frequencies are chosen based on their ability to resonate for a given mini-structure and can be different for different smartphone case designs. The challenge will propagate through the physical device, transformed by the joint structure of the user hand, smartphone, and mini-structures. During propagation, the user touches a mini-structure, which embeds the finger physiology information and transforms the challenge. The transformation creates a user-specific response that is measured by built-in microphones.

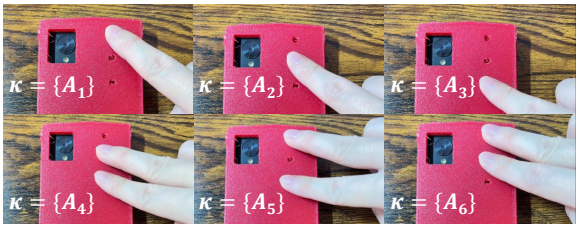


Figure 7: Example of pass code input using BioCase. The smartphone is face down for visual clarity of the finger-mini-structure interactions.

Response Pre-Processing. We isolate the transmitted challenge from background noise and filter out any sources of interference. A bandpass filter is applied to remove irrelevant frequency bands. Once segmented and filtered, we confirm the integrity of the response by verifying the frequency of each received chirp. Only the user smartphone is aware of the challenge frequency pattern; fabricated signals from attackers will be rejected as they cannot know the randomly generated order of the challenge.

Feature Extraction. User finger features are extracted and used to construct a preliminary feature vector. Simultaneously, the frequencies of each received chirp are inspected to determine if a corresponding mini-structure was sealed by the finger touch to diminish the resonant frequency. The frequency response can be controlled by the user based on the mini-structure they choose to touch. We detect this choice by computing the degree of transformation to each chirp frequency; larger differences implies a specific mini-structure was touched. This ensures that adversaries cannot be mistaken for the user, even with similar physical characteristics, without knowing the user's mini-structure interactions.

Profile Decision Making. We determine if the response matches the profile of the legitimate user by comparing the received response with saved feature profiles. Our system utilizes a support vector machine (SVM)-based classifier to perform this comparison. A pre-installed database of anonymized challenge-responses act as negative labels. The verification decision can be used by other applications to selectively display private information, such as text messages, or unlock the smartphone device itself.

4.2 Encoding Biometric-Hybrid Signatures

Biometric credentials may be stolen or impersonated by adversaries with no recourse for the legitimate user. Therefore, it is also necessary to challenge the biometric information presented to ensure authenticity. We achieve this by controlling not only the acoustic signal, but the structure itself as well. We devise a code $\kappa = \{A_1, ..., A_N\}$ where each member corresponds to an acoustic response produced by user interaction with a mini-structure in the smartphone case. Each member of the alphabet is a response derived from Eq. 2. The size of the code alphabet N is determined by the number of mini-structures n and the number of possible user interactions with the mini-structures m, generalized by the following equation:

$$N = n + C_m^n + 1, \tag{4}$$

We define C_m^n as the unordered combinations of length m derived from the set of n elements or $C_m^n = \frac{n!}{m!(n-m)!}$, where $m < n$. For example, if $n = 3$ and we allow the user to interact with at most 2 mini-structures simultaneously (to avoid the dilemma of Eq. 3), then the number of possible user interactions and alphabet size is $N = 3 + C_2^3 + 1 = 7$. Note that our alphabet can also match or exceed PIN code complexity (e.g., a 4th mini-structure $n = 4$ yields a code size $N = 11$ under the same assumptions). Unlike PIN codes, however, our system does not require an arbitrarily large number of code combinations to guarantee security. User physiology and behavior is also recognized based on where and how they touch the smartphone case. The user can therefore prove their identity by inputting a touch-based pass code as shown in Figure 7. Pass code P is a secret combination of finger(s) and mini-structure(s) that the user may choose during enrollment.

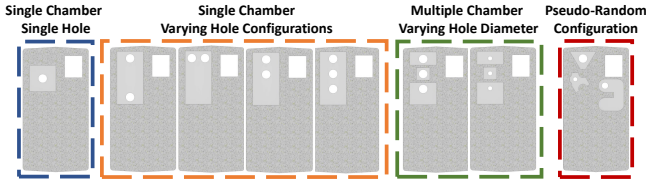


Figure 8: Iterative prototyping of BioCase designs.

4.3 Mini-structure Design

Mini-structures along the sound propagation path influences the frequency response of the pass code [20]. In order to develop a smartphone case capable of differentiating users through finger touches, we iterated our case design over dozens of single-purpose prototypes and recruited volunteers to better understand how to control and isolate the acoustic responses. This data-driven prototyping process enabled us to achieve a more optimized smartphone case design that alters the acoustic properties of structure-borne sound propagating through a controlled pathway.

The evolution of our prototyping process can be seen in Figure 8. Based on Eq. 1, the design of the mini-structure is a relatively open-ended problem so long as the geometry constrains the structure-borne sound propagation to diffract with constructive interference at the desired resonant frequency. We find that using multiple miniature chambers are more likely to produce distinct frequency responses due to their more dissimilar geometries. For modeling simplicity, we design these structures to be rectangular prisms. Our mini-structures are positioned vertically as directly in the propagation path as possible. We find that smaller holes allow users to more consistently seal the mini-structures and therefore produce repeatable patterns in the challenge-response. We also determined that multiple holes to the same mini-structure will diminish the ability of each hole to alter the acoustic properties. Hence, we feature only a single hole to each mini-structure. We also develop a pseudo-random case that does not follow any specific design philosophy to verify these our observations through experimental comparison. Note that the locations of our mini-structures are flexible. We envision that BioCase variants with more ergonomic considerations can allow users to touch virtually anywhere for authentication.

4.4 Challenge Signal Design

A chirp signal is transmitted by the built-in speakers of the smartphone to stimulate a response in the user hand while operating the

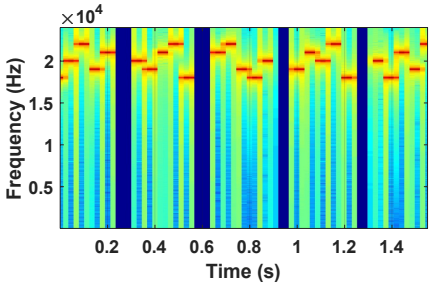


Figure 9: Example of five challenge signals. Frequency sweeps are random each transmission.

device [21]. Both time-domain and frequency-domain parameters must be considered for the chirp design.

Time-Domain. Unlike most acoustic works, we consider structure-borne sound to be the target signal and airborne sound to be noise. Separating the two is possible by controlling the time duration of the chirp transmission; sound propagation is faster in solids than air. By keeping chirps short and buffer space between transmissions sufficiently large, the arrival of airborne noise can be staggered such that it does not coincide with structure-borne sound.

Frequency-Domain. To minimize disturbances, the signal should be inaudible while also being robust to interference. The ultrasonic frequency band ($\sim$ $16kHz$ and above) is both inaudible to most adults and beyond the frequencies of common daily noise sources, simplifying filtering. However, most commercial smartphone microphones are limited in sampling rate to $\sim$ $48kHz$, bounding the upper frequency limit to $24kHz$ according to Nyquist-Shannon Theorem.

Accordingly, we design a set of candidate chirp signals for transmission. Each chirp sweeps one of five singular inaudible frequencies (i.e., $18kHz$ to $22kHz$ in steps of $1kHz$). The duration lasts only 25ms to aid the separation of structure-borne sound from multipath reflections. From our experiments, we observed that commercial smartphone speakers may struggle to consistently sweep such high frequencies in a short time period. Therefore, we transmit multiple chirps in succession to ensure a viable response can be captured by the microphones. During the acoustic challenge, each chirp of the transmission $S(\omega)$ will be embedded with biometric-hybrid signatures from the user's finger geometry and interaction with the mini-structures. Biometric data may be vulnerable to impersonation if static (e.g., still image of the face). To avoid this problem, the frequency of each chirp in the challenge signal is randomly selected at transmission, seen in Figure 9. Attackers may eavesdrop past challenge-responses but cannot reuse them since future challenges will differ.

The receiving microphone, knowing the design of $S(\omega)$, can locate and verify the authenticity of the challenge by applying cross-correlation with the response. We compute correlation through the function $\sum_{\omega=0}^{\infty} S^*(\omega)R(\omega+t)$ where $S^*(\omega)$ is the complex conjugate of our transmission and $R(\omega+t)$ is the recorded propagation sequence time shifted by some unknown delay t. Both $S^*(\omega)$ and $R(\omega+t)$ are normalized to compensate for amplitude differences in the differing sound propagation. By locating the index with the highest correlation to our transmission, we can determine the start point of our signal. Environmental reflections arrive at a much slower speed, thus we can safely remove interference by keeping our transmission short (i.e. milliseconds long), calculating the signal endpoint based on sequence length, and segmenting audio at this point.

5 IMPLEMENTATION

5.1 Feature Extraction

We construct a user finger feature profile that integrates both finger physiology and user knowledge (i.e., mini-structure pass code). A series of candidate features are identified through empirical tests similar to Section 2.3.

Time-domain Features. In the time domain, we extract statistics including mean, standard deviation, maximum, minimum, range, kurtosis and skewness. We also estimate the response's distribution by calculating second quantile, third quantile, fourth quantile, and signal dispersion. We examine peak change by deriving the index position of the data point that deviates most significantly from the statistical average.

Frequency-domain Features. In the frequency domain, we apply Fast Fourier Transformation (FFT) to the response and derive 256 spectral points to capture the unique characteristics of the user's holding hand. We also derive the acoustic features from the received sound using the MFCC. Our response signal is first processed using a first-order FIR filter with a pre-emphasis coefficient value of 0.97. The filtered signal is then subjected to the Discrete Cosine Transform to produce our MFCC features consisting of 13 filter bank coefficients. Intuitively, touching mini-structures will suppress some, but not all, resonant frequencies.

Pass Code Detection. The challenge signal $S(\omega)$ contains a sequence of chirps assembled from a pool of candidate chirps of varying frequencies. We compute the differences between $S(\omega)$ to $R(\omega)$ as $d_i = \sqrt{|S(\omega) - R(\omega)|^2}$ and use d_i to scale the user hand feature vector. Minor differences already exist between code words $\kappa(A_i)$ and $\kappa(A_j)$ when performed by the same user due to changes in finger position and mini-structure. Feature vector scaling thus magnifies pre-existing differences to aid in distinguishing user.

5.2 User Profiling

We develop learning-based algorithms to recognize the unique characteristics of the user from the acoustic challenge-response. This information is used to verify whether the finger and code input match the known profile of the legitimate user. An existing data set of anonymized user profiles is included to serve as negative labels when classifying the user during authentication attempts. For each chirp component of $R(\omega)$, our algorithm utilizes the prediction probabilities returned by the classifier as a confidence level and applies a threshold-based method to examine the classification results. If the confidence level of the classification is above the threshold, the user is authenticated and otherwise rejected. A majority vote based on the confidence of all chirp signals is conducted to confirm the decision. This decision can be used for multiple applications such as locking devices when an unknown user attempts to gain access. We choose a Support Vector Machine (SVM)-based architecture with 10-fold cross-validation for our classifier due to its efficacy at processing features with high dimensionality.

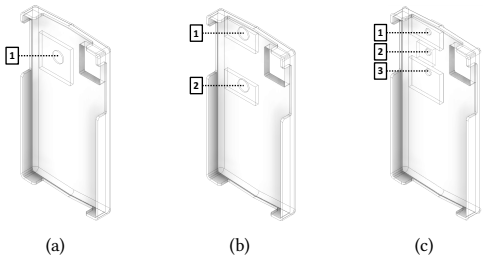


Figure 10: Isometric view of BioCase prototypes. Mini-structures are designated a number for reference.

6 EVALUATION

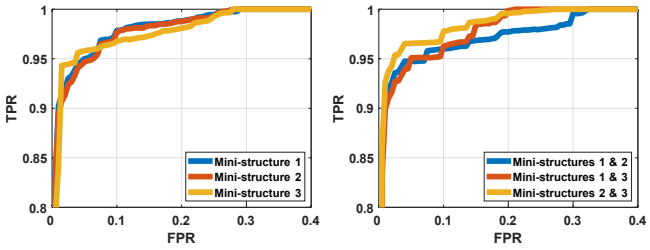
6.1 Data Collection

Hardware. A prototype app of BioCase was developed for Android platforms. Our smartphone cases and mini-structures are designed in Autodesk Inventor and exported in STL format for slicing in commercial 3D printers. We fabricate physical prototypes using Prusa I3MK3S printers and choose TPU plastic filament for it's low cost and usage in commercial smartphone cases. An isometric view of our designs, based from our study in Section 4.3, are shown in Figure 10. Our cases were designed to fit the *Nexus 5* and *Galaxy Note 5* smartphones. We assume the user may use up to two fingers to touch two mini-structures simultaneously, yielding a code alphabet size of $N = 7$. We also produce alternate designs of our prototypes for different mini-structure designs and filament material to study the impact on sound propagation.

Scenarios. We recruit 46 volunteers, 40 males and 6 females ranging from ages 18-60, to participate in our study with the approval of our IRB. We evaluate our system by conducting the experiments in a typical office with an average 35dB background noise level. We simulate additional environments by playing corresponding background noise at appropriate sound pressures (e.g., construction noise at 80db). Tested scenarios include using different fingers with BioCase, simulating different attacks, and studying impact factors such as noise and motion. All experiments are conducted using the Nexus 5 in a typical office environment without strong noise unless otherwise noted. For each use case scenario and each participant, we collect 80 acoustic challenge-responses of finger touch information or over 3,600 user samples. To ensure diversity in user samples, the participants were asked to pick and put down the phone several times. Participants are also encouraged to hold the smartphone naturally based on their own habits (e.g., tightness of the grip, preferred hand). In our evaluation, we iterate over all participants where we treat each participant as the legitimate user of the smartphone and all remaining participants as non-legitimate users or attackers. We do not consider our acoustic responses to be personally revealing information but will maintain this data privately as a precaution.

6.2 Performance Analyses

We describe accuracy of our system by evaluating the relation between True Positive (TP), False Positive (FP), and False Negative (FN) rates. We compare TP and FP through receiver operating characteristic (ROC) curves to measure general performance and compare FP and FN through equal error rates (EER) to quantify balance between security and convenience. For our purposes, higher TP means higher probability the legitimate user is recognized, higher FP means higher probability for non-legitimate users to be mistaken for the user, and higher FN means higher probability that the user is not recognized. The ideal system has a simultaneous 100% TP rate, 0% FP rate, and 0% FN rate (i.e., the user is always recognized, never rejected, and attackers are always rejected). We evaluate BioCase by answering four main questions regarding usability and security aspects.



(a) Single mini-structure interactions using index finger.

(b) Dual mini-structure intereaction using the index and middle fingers.

Figure 11: Performance for finger-only interactions.

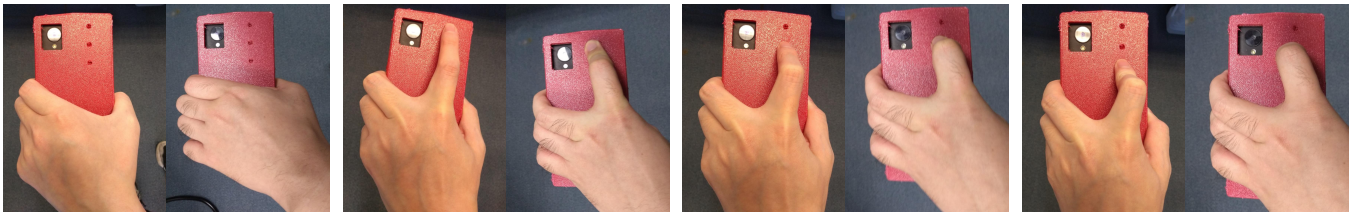
6.2.1 *How well does BioCase authenticate legitimate users?*

Finger-Only Interaction. We study scenarios where the smartphone is not held by the user, such as when resting on a table. We equip our cases to the smartphone and position the mini-structures face up. Each user is asked to attempt every pass code input possible for the given case design (i.e., the case in Figure 10(a) has the fewest while Figure 10(c) has the most). The experiment is repeated for each case from Figure 10. For all 46 participants, we let each act as a legitimate user of the smartphone device while others act as non-legitimate users. Figure 11 shows we can achieve an average 94.4% TP rate for a simultaneous 5% FP rate using biometric-hybrid signatures gathered by BioCase. In particular, we find single finger interactions to be slightly more effective at distinguishing users (95.6% TP) than two fingers (93.2%). Meanwhile, the EER is under 4% for both interactions. Our results suggest finger physiology is recognized as 46 users touching the same mini-structure can still be distinguished.

Hand and Finger Interaction. We evaluate the ability of the privacy protection system to distinguish user responses while operating the smartphone in their hand. We study each user's response single mini-structure interactions. We also compare with no mini-structure interactions (i.e., phone simply held in hand) for a baseline comparison. Figure 12(a)-12(d) depict examples of users interacting with BioCase, picturing the back of the smartphone to show finger-mini-structure interactions. Figure 12(e)-12(h) shows the average performance when authenticating all participants for each possible code word. Specifically, we can achieve an average 95.1% TP rate for a simultaneous 5% FP rate and an EER of 3.79% on our case with 3 mini-structures. Our performance is comparable to that of palm biometric sensing methods. However, our finger-based biometric-hybrid signature also adds robustness in the form of multiple viable pass codes (i.e., mini-structure interactions).

6.2.2 *How secure is BioCase when under various attacks?*

Impersonation Attack. Our attack setup is similar to EchoLock [42] where we involve a subset of 10 participants. One participant

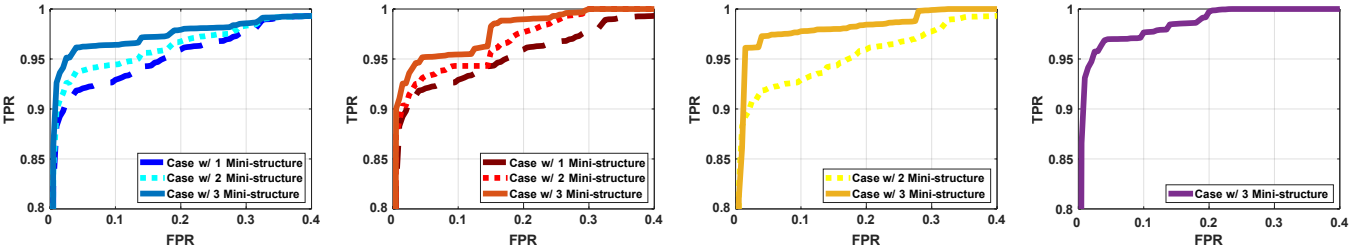


(a) Example of two users with no mini-structure interaction.

(b) Example of two users pressing mini-structure 1.

(c) Example of two users pressing mini-structure 2.

(d) Example of two users pressing mini-structure 3.



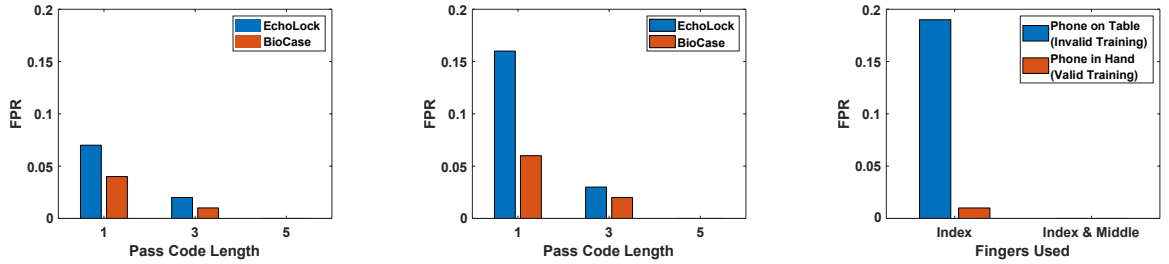
(e) Performance when interacting with no mini-structures.

(f) Performance when interacting with mini-structure 1.

(g) Performance when interacting with mini-structure 2.

(h) Performance when interacting with mini-structure 3.

Figure 12: Authentication performance when distinguishing users based on the pass code input using one finger.



(a) Blind impersonation attack.

(b) Informed impersonation attack.

(c) Puppet attack.

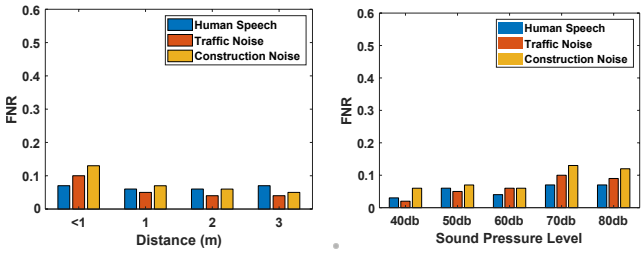
Figure 13: Performance under different attack types that may be launched against BioCase. We consider scenarios where the attacker attempts to gain unauthorized access with varying degrees of knowledge about the user.

acts as our victim and the remainder act as attackers in round robin order, meaning dozens of user pairs are tested. We consider scenarios where the user may prioritize security over convenience and choose multiple code inputs in sequence to verify themselves, akin to using longer passwords. We evaluate blind attacks, where attackers have no knowledge of the user or code, and informed attacks, where attackers are explicitly told the pass code and allowed to observe the user for 30 seconds to understand the victim's finger-mini-structure interactions. For all scenarios, the attackers are given 10 attempts to imitate the user. For state-of-the-art comparisons, we develop an implementation of EchoLock [42] that does not use smartphone cases and repeat the experiments.

Figure 13(a) and 13(b) indicate that impersonation attacks are not viable, showing only a 6% rate at being falsely identified as the legitimate user in the best case impersonation scenario of a short pass code and informed attacker. EER can be maintained under 5% for all tested pass codes. We achieve substantial improvement over existing work, reducing false positives by nearly 10% in the informed impersonation scenario. We observed that the majority of false positives were consistently caused by the same attacker-victim pair, suggesting that physical hand similarity may be more crucial to successful impersonation. However, attackers cannot control how similar their hands are to potential victims, nor always know the pass code in advance. This suggests that user physiology is captured in our responses as different users pressing the same location are not easily mistaken for one another.

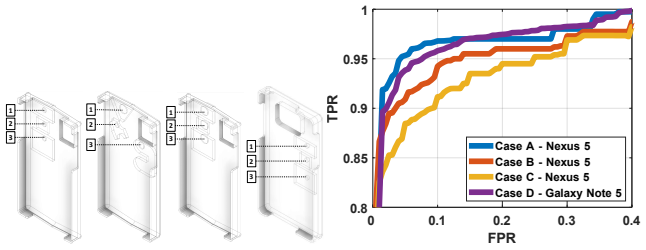
Replay Attack. We consider attacks using eavesdropped signals and synthesized signals. To obtain eavesdropped signals, we seat the victim user at a desk while they operate our privacy protection system. A separate mobile device positioned 0.2m from the victim acts as a malicious sensor, listening for the acoustic response. Many devices can activate microphones without obvious indicators, making this attack strategy highly plausible. For the synthesized signal, we generate 10 challenge signals for the attacker to replay. Note that the random seed is different for each synthesis. We recruit 10 of our participants to act as 9 victims and 1 attacker. The profiles of the remaining participants are used as negative labels during identification. We exclude attacker data from training to simulate an attack by an unknown user.

BioCase is able to correctly block the attacker when attempting to replay eavesdropped and synthesized signals, resulting in 0% EER and 100% TP rates for both scenarios. For the synthesized replay



(a) Varying distance to noise source. (b) Varying sound pressure level.

Figure 14: Performance in noisy environments with common sources of acoustic interference.



(a) (b) (c) (d) (e)

Figure 15: Performance comparison for alternate case designs featuring three mini-structures.

scenario, we also ensured at least one instance where the genuine challenge matched the synthesized challenge. Both eavesdropped and synthesized signals fail to be accepted by BioCase as the replayed signal loses structural properties when attenuating through the air from attacking speaker to victim microphone.

Puppet Attack. We simulate scenarios where the user may have their fingers interact with BioCase without their consent (e.g., asleep). We recruit 9 participants to act as victims and one to act as the attacker. We place the smartphone on a flat table surface with the mini-structures face up. For each victim user, the attacker moves the victim fingers to press against the mini-structure hole. To simulate the best case scenario for the attacker, we assume the attacker is informed (i.e., knows where to press). During the attack, the victim keeps their hand limp. The puppet response is compared to profiles trained when the smartphone is lying on a table (invalid training, best scenario for the attacker) and when the smartphone is held in hand (valid training, most common scenario).

The results of Figure 13(c) suggest that puppet attacks are unlikely to succeed in most conditions. Notably, we find interactions using two fingers (index and middle fingers) to be too complex for attackers to manipulate, resulting in 0% FP in all tested conditions. Meanwhile, single finger interactions were only possible when the user profile was trained in a similar condition to the attack scenario (i.e., phone lies on table). However, we consider this training style invalid as most users will not enroll their profile without holding the device. Additionally, we assumed the attacker is informed and already knows the correct finger interaction. This implies our responses capture user behavior as well since the attacker cannot pass as the victim, even when possessing the legitimate finger physiological information and pass code. For uninformed attackers, brute forcing the attack becomes much more complex due to the possible finger-structure combinations. Brute force can be defended against by imposing a limit on number of attempts permitted.

6.2.3 *How stable is BioCase under different impact factors?*

Noise Pressure Level and Distance. In order to test the performance of BioCase in various noisy environments, we play three categories of noise at different sound pressure levels and distances from the user using an external speaker. When controlling distance to the noise source, we calibrate our noise pressure level to 80dB as measured by the smartphone at 0.5m and progressively test performance at longer distances. When controlling noise pressure level, we calibrate the noise to the desired dB level as measured by the smartphone at 1m distance. Noise sources include human voices, traffic sounds such as car honks, and construction noise such as jackhammers. Figure 14 indicates that the lower frequencies of most audible noise (human speech, traffic) have limited impact on our ultrasonic frequency band. FN rates are consistently under 10%, regardless of distance. The higher frequencies of drilling and jackhammers were more challenging, having the highest FN rate. However, most participants found this noise source annoying. We believe most users will stop using their smartphones in order to distance themselves from this noise, mitigating the problem.

Different BioCase and Smartphone Designs. We design and manufacture variations of our BioCase mini-structures for use in experimentation. We first study variations for our three-mini-structure design on the Nexus 5. Upon determining the best performing case, we adapt the design for other smartphone models. We depict examples of these cases in Figure 15 and compare their performance at distinguishing user challenge-responses. As seen in Figure 15(e), our primary design, denoted as Case A, has the highest TP rate of 94% for a 5% FP rate. Meanwhile, the EER is 3.79%. Notably, the pseudo-random mini-structures of Case B performed slightly better (91% TP rate) than Case C, our variation of Case A with larger hole diameters (89% TP rate) for the same FP rate. Although the pseudo-randomness of Case B is not intentionally optimized, the drastic differences in mini-structure configuration may be distinct enough to produce unique responses for each mini-structure. Case C has considerably lower accuracy than Case A, possibly due to the larger holes being more difficult for users to seal with fingers consistently.

Because of its higher performance, we replicated the mini-structure design of Case A in a BioCase prototype for the Galaxy Note 5. We found performance on the Galaxy Note 5 to be competitive with our Nexus 5 device, outperforming the Case B and C designs while achieving 93% TP for a 5% FP rate and a 4.1% EER. While we observed success in adapting BioCase from the Nexus 5 to the Galaxy Note 5, the larger size and different sensor placements (e.g., camera) of the Galaxy Note 5 may also enable new design choices not possible on the Nexus 5 that could further improve performance.

Different Fingers. We study performance when using different fingers to interact with BioCase. We leave the smartphone on a table with the case mini-structures face up and allow 10 users to input their pass codes. We also ask the users to use each of their five fingers when interacting with the different mini-structures. Figure 16 shows we can achieve an average 92.1% TP rate for a simultaneous 5% FP rate across all fingers. EER can be maintained under 10% for all fingers. Note that performance can vary depending on

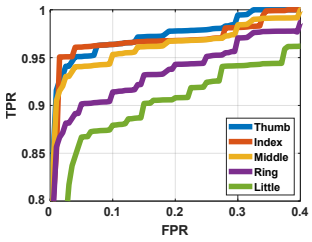


Figure 16: Performance for different fingers.

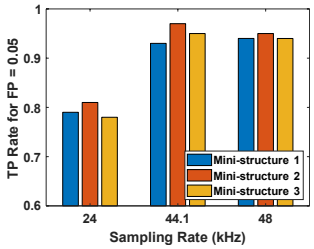


Figure 17: Performance for sampling rates.

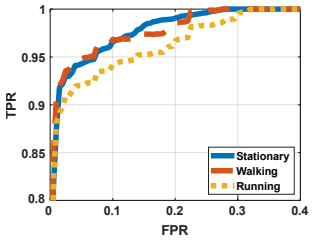


Figure 18: Performance for user motions.

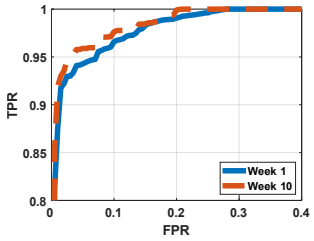


Figure 19: Performance after extensive time passed.

the specific finger used. During ordinary smartphone usage, users primarily use their index and middle fingers, which were among the highest performing fingers. While thumbs also perform well, users typically rest their thumbs on the front of the smartphone rather than the back, making them less intuitive to use. Ring and little finger interactions are noticeably more difficult to recognize, possibly due to most users having lower dexterity in these fingers. We consider this comparable to fingerprint scanning as users rarely choose to scan ring or little fingers.

Sampling Rate. The sampling rate determines the quality of the received signal. Signals received by high sampling rate can provide more information at the expense of storage and power consumption overhead. Therefore, we conduct experiments to examine the performance of our system by lowering our preferred $48kHz$ sampling rate to $44.1kHz$ and $24kHz$. From Figure 17, it can be observed that our system can maintain stable 94% TP rate for the comparable $44.1kHz$ whereas halving the sampling rate to $24kHz$ noticeably lowers performance to 79% TP. Due to the high frequencies utilized by our acoustic challenge-response, lowering the sampling rate can cause aliasing in our measurements. However, if audibility is not a concern, lower acoustic frequencies could be utilized and allow lower sampling rates.

User Motion. Users may operate their smartphones in mobile settings such as while walking or during exercise. We therefore verify that the acoustic sensing process of BioCase is not disrupted by such motions. We consider two actions specifically; walking and running. To create the most natural experimental conditions, we only ask the participants to walk or run at a comfortable pace for several minutes. No further instruction on body motions is given. Therefore, the precise behavior varies slightly between users, but all participants exhibited ordinary behaviors (i.e., swaying arms, rhythmic steps). We are able to generalize the motions as approximately one step per second for walking (comparable to daily routine walking) and two to three steps per second for running (comparable

(a) (b) (c)

Figure 20: Examples of commercial smartphone cases studied. Finger touch locations are highlighted.

to light jogging exercise). The results of Figure 18 indicate these actions do not have significant affects on BioCase performance. Only minor deterioration was observed for running, showing decreases in TP by 3% and EER increases by 2% on average for 10 participants. While higher running speeds are possible and could negatively impact performance, we believe few users simultaneously use their phones during such actions due to safety risks. Furthermore, attacks are difficult to launch against a running person, reducing security risks.

Time Variance. We verify that our acoustic challenge-response is sufficiently stable to authenticate the user even after significant time passes. We invited 10 of our participants to return and interact with BioCase 10 weeks after their initial participation. Figure 19 shows that BioCase can still successfully authenticate individuals, achieving an average 95.2% TP for a simultaneous 5% FP or lower and an EER of 3.75%. This result is actually slightly better than initial averages recorded immediately after enrollment, possibly due to users already being familiar with BioCase. We note this performance can be further improved by periodically sampling new challenge-responses from the user.

7 DISCUSSION

System Overhead. We find that the memory required for the challenge-response signals is ~2.5MB while the authentication model itself is only 15KB. The total size of the BioCase app is 14MB, which is less than most commercial smartphone apps. For instance, the Spotify app officially recommends 1GB of memory [33]. In addition, we estimate the battery costs of BioCase by calculating the number of authentication attempts necessary to deplete 1% of total battery based on Android app settings. We find that our method consumes negligible power (i.e., 0.036% battery capacity per authentication) based on tests across 10 participants, 20 authentication attempts per person, and two smartphones.

System Generalizability. We test our system with commercial smartphone cases, as shown in Figure 20. Each case has distinct physical traits (e.g., built-in stands, decorative indents, shock absorbers). We found that the average performance of distinguishing 10 users is around 83% TP rate. This is an encouraging sign, as performance is still moderately high for smartphone cases not designed with structure-borne sound in mind. This observation indicates that mini-structure design is highly influential in the quality of acoustic responses produced by BioCase. With minor modifications, manufacturers can integrate compatible mini-structures and enable an additional layer of security at a low cost.

System Usability. Our system is highly usable with low latency. We only transmit 5 chirp signals, and each lasts only 25ms. When incorporating computation time (variable by hardware), we can achieve latency under 500ms, similar to existing authentication methods. In addition, BioCase is more accessible and lower in cost than specialized sensors like fingerprint scanners [19, 37]. We allow the user to perform a single finger touch to authenticate themselves, which is similar to traditional fingerprint scanning in terms of user-friendliness. Longer challenges or touch sequences may improve performance at the expense of longer wait times. We leave a detailed study of this for our future work.

Additional Case Design Factors. This is a pioneering work leading to the development of more ergonomic, user-friendly BioCase designs. To this end, we plan to explore more mini-structure factors (e.g., distance, shape, size, and positions) in our future work. We also intend to examine the impact of different case materials, which may affect sound propagation. TPU is a flexible plastic and one of the most popular materials for building off-the-shelf smartphone cases. The cost to manufacture one of our prototypes with TPU is competitive with existing commercial case prices (around 5 to 10 USD). We note that cosmetic scratches have minimal impact on the structure of the TPU case. Although severe trauma or extensive aging could damage the structure and change the sound propagation behavior, the user is unlikely to continue using the same case in such scenarios. Such damaged smartphone cases are easy to replace, unlike sensors. Users can also quickly enroll a new profile through a one-time training process, similar to providing their fingerprints on a new phone. In addition to TPU material, our experiments also include an alternative Nylon 12 polymer case, which find negligible differences in performance. We will consider other materials in the future.

Acoustic Signals and Sensors. The high frequencies utilized by BioCase are generally benign and difficult for adults to hear, but could be perceived by younger individuals [30]. We keep transmission short to reduce risks. We interviewed all participants after experiments to understand their experience and found none were able to hear the transmissions. Our system requires only a single speaker-microphone pair. In this work, we leverage the speaker and microphone located at the top and bottom of the smartphone, which is a common setting in most of off-the-shelf smartphones. There may be other configurations (e.g., speakers and microphones on the front/back) that could affect our system. We leave further study of these configurations for future work.

Changes to User Hand. We studied the impact on performance when the user's hand is in atypical conditions. We invited five participants to interact with BioCase in scenarios such as dampening hands with sweat or wearing thin wool gloves. Our system is robust even under such conditions, with no significant decrease in authentication performance. We also consider variances in users' hand interactions. Inconsistent touch pressures (i.e., stronger or weaker) may alter the amount of user information the system may capture. Similarly, different phone holding styles will change the degree of physical contact with BioCase, which may have adverse effects on performance. During data collection, participants were asked to hold the device naturally. However, a more controlled study can be conducted in future work to better understand the impact of these factors.

8 RELATED WORK

Traditional privacy protection favors text-based passwords [27], graph-based swiping patterns [38], and image-based picture recognition [7, 35], which requires memorization of long inputs (e.g., random characters or pictures) and verifies possession of keys rather than user identity. To ensure credentials are submitted by the user, biometric-based schemes (e.g., face recognition [5, 8, 28], fingerprinting [40, 41, 43], iris scanning [2, 17, 31], retina patterns [26, 29]) use physical traits of the users for authentication. However, these methods require dedicated hardware components for precise measurements, limiting the devices they can be deployed on.

Fortunately, all smartphones possess microphones and speakers, enabling acoustic sensing. Most acoustic-based user authentication leverage voice characteristics for speaker recognition [3, 11, 14]. However, these features are vulnerable to impersonation or spoofing attacks [16]. To minimize such risks, some works integrate biometric and behavioral traits. LipPass [24] leverage build-in audio devices on smartphones to extract unique behavioral characteristics of speaking lips for user authentication. BreathPrint [4] uses the audio signatures associated with breathing to identify users. Other strategies include challenge-responses based on bodily reflections. EchoPrint [45] uses inaudible acoustic signals to scan the user's face and authenticate features of the reflecting signals. However, this still requires active user participation to aim speakers towards the face. EchoLock [42] verifies the user based on palm biometrics (i.e., hand geometry, holding strengths and holding styles) using structure-borne sound propagation, which can be threatened by sophisticated forgery attacks [9, 25]. EchoHand [39] combines acoustic sensing with hand image recognition to improve robustness, but raises hardware costs by requiring cameras. SonicPrint [32] detects the sound of finger swipes to authenticate the user, which requires more user effort than our simple finger touches. BioCase is most similar to the latter category of techniques. Unlike prior works, we capture both palm and finger information and demonstrate enhanced security and robustness due to our fusion of biometric-based and knowledge-based credentials. We achieve a similar performance while testing more participants and devices. We also more study how the physical structures of smartphones can enhance authentication performance. We develop low-cost smartphone cases capable of integrating biometric sensing with knowledge-based verification via finger touch inputs.

9 CONCLUSION

We presented BioCase, an acoustic-based smartphone privacy protection system that authenticate users via finger touches. An inaudible acoustic challenge signal is used to stimulate the smartphone structure and produce a user finger response. By embedding mini-structures of specific shape and design into smartphone cases, we can control characteristics of structure-borne sound propagation. Mini-structure location imparts spatial diversity in the response, revealing where the user touches. Mini-structure dimensions control resonant frequency, imparting frequency diversity in the response and enabling secret code inputs by the user. This enables our fusion of biometric and knowledge-based credentials into a biometric-hybrid signature for verification. We demonstrate that BioCase

provides robustness against multiple attacks. Evaluations of multiple use case scenarios indicate we can identify users with over 94% TP rate and under 5% FP rate.

ACKNOWLEDGMENT

This work was partially supported by the National Science Foundation Grant CNS2120396.

REFERENCES

[1] Abdulrahman Mohammed Obaid Almheiri, Shanaihi Sanjay Patel, and Bhoopesh Kumar Sharma. 2022. A Conceptual Study of Forgery of 3D Fingerprints and Its Threat to Biometric Security Systems. *Journal of Positive School Psychology* (2022), 4453–4462.

[2] Silvio Barra, Andrea Casanova, Fabio Narducci, and Stefano Ricciardi. 2015. Ubiquitous iris recognition by means of mobile devices. *Pattern Recognition Letters* 57 (2015), 66–73.

[3] Joseph P Campbell. 1997. Speaker recognition: A tutorial. *Proc. IEEE* 85, 9 (1997), 1437–1462.

[4] Jagmohan Chauhan, Yining Hu, Suranga Seneviratne, Archan Misra, Aruna Seneviratne, and Youngki Lee. 2017. BreathPrint: Breathing acoustics-based user authentication. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. 278–291.

[5] Shaxun Chen, Amit Pande, and Prasant Mohapatra. 2014. Sensor-assisted facial recognition: an enhanced biometric authentication system for smartphones. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. 109–122.

[6] Ivan Cherapau, Ildar Muslukhov, Nalin Asanka, and Konstantin Beznosov. 2015. On the Impact of Touch ID on iPhone Passcodes. In *Symposium on Usable Privacy and Security (SOUPS)*. 257–276.

[7] Sonia Chiasson, Paul C van Oorschot, and Robert Biddle. 2007. Graphical password authentication using cued click points. In *Computer Security–ESORICS 2007*. Springer, 359–374.

[8] Mohammed E. Fathy, Vishal M. Patel, and Rama Chellappa. 2015. Face-based Active Authentication on mobile devices. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*. IEEE.

[9] Vishu Gupta, Masakatsu Nishigaki, and Tetsushi Ohki. 2019. Unsupervised Biometric Anti-spoofing using Generative Adversarial Networks. *International Journal of Informatics Society* 11, 1 (2019), 5.

[10] Ke Han, Zizheng Wang, and Zilong Chen. 2018. Fingerprint Image Enhancement Method based on Adaptive Median Filter. In *2018 24th Asia-Pacific Conference on Communications (APCC)*. 40–44. https://doi.org/10.1109/APCC.2018.8633498

[11] Matthieu Hébert. 2008. Text-dependent speaker recognition. In *Springer handbook of speech processing*. Springer, 743–762.

[12] Long Huang and Chen Wang. 2022. PCR-Auth: Solving Authentication Puzzle Challenge with Encoded Palm Contact Response. In *2022 IEEE Symposium on Security and Privacy (SP)*. 1034–1048. https://doi.org/10.1109/SP46214.2022.9833564

[13] Apple IOS. 2019. Face ID. https://support.apple.com/en-us/HT208108.

[14] Apple IOS. 2019. Siri. https://www.apple.com/ios/siri/.

[15] Andy Kiersz. 2014. REVEALED: Here's Who Uses iPhone Cases And Why. https://www.businessinsider.com/iphone-case-survey-2014-7.

[16] Tomi Kinnunen, Bingjun Zhang, Jia Zhu, and Ye Wang. 2007. Speaker verification with adaptive spectral subband centroids. In *International Conference on Biometrics*. Springer, 58–66.

[17] Ajay Kumar and Arun Passi. 2010. Comparison and combination of iris matchers for reliable personal authentication. *Pattern recognition* 43, 3 (2010), 1016–1026.

[18] Alexander Kunst. 2019. Protective case usage among U.S. smartphone owners 2017. https://www.statista.com/statistics/368627/us-protective-case-usage-among-smartphone-owners/.

[19] Alexander Kunst. 2019. Protective case usage among U.S. smartphone owners 2017. https://www.statista.com/statistics/368627/us-protective-case-usage-among-smartphone-owners/.

[20] Gierad Laput, Eric Brockmeyer, Scott E Hudson, and Chris Harrison. 2015. Acoustruments: Passive, acoustically-driven, interactive controls for handheld devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2161–2170.

[21] Jingjie Li, Kassem Fawaz, and Younghyun Kim. 2019. Velody: Nonlinear vibration challenge-response for resilient user authentication. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. 1201–1213.

[22] Jie Li, Wenyao Yang, Yidan Shi, and Zedong Nie. 2020. Design of Finger Vein Recognition SOC Based on FPGA. In *2020 7th International Conference on Information Science and Control Engineering (ICISCE)*. 2385–2389. https://doi.org/10.1109/ICISCE50968.2020.00468

[23] Yuxi Liu and Dimitrios Hatzinakos. 2014. Human acoustic fingerprints: A novel biometric modality for mobile security. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 3784–3788.

[24] Li Lu, Jiadi Yu, Yingying Chen, Hongbo Liu, Yanmin Zhu, Yunfei Liu, and Minglu Li. 2018. Lippass: Lip reading-based user authentication on smartphones leveraging acoustic signals. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 1466–1474.

[25] Deborah Lupton. 2015. Fabricated data bodies: Reflections on 3D printed digital body objects in medical and health domains. *Social Theory & Health* 13 (2015), 99–115.

[26] Cástor Mariño, Manuel G Penedo, Marta Penas, María J Carreira, and F Gonzalez. 2006. Personal authentication using digital retinal images. *Pattern Analysis and Applications* 9, 1 (2006), 21–33.

[27] Robert Morris and Ken Thompson. 1979. Password security: A case history. *Commun. ACM* 22, 11 (1979), 594–597.

[28] P Jonathon Phillips, Patrick J Flynn, Todd Scruggs, Kevin W Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. 2005. Overview of the face recognition grand challenge. In *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, Vol. 1. IEEE, 947–954.

[29] Abhinand Poosarala. 2018. Uniform classifier for biometric ear and retina authentication using smartphone application. In *Proceedings of the 2nd International Conference on Vision, Image and Signal Processing*. 1–5.

[30] Dale Purves and Stephen Mark Williams. 2001. *Neuroscience. 2nd edition*. Sinauer Associates 2001. http://lib.ugent.be/catalog/ebk01:3450000000002013

[31] Kiran B Raja, Ramachandra Raghavendra, Martin Stokkenes, and Christoph Busch. 2015. Multi-modal authentication system for smartphones using face, iris and periocular. In *Biometrics (ICB), 2015 International Conference on*. IEEE, 143–150.

[32] Aditya Singh Rathore, Weijin Zhu, Afee Daiyan, Chenhan Xu, Kun Wang, Feng Lin, Kui Ren, and Wenyao Xu. 2020. SonicPrint: A Generally Adoptable and Secure Fingerprint Biometrics in Smart Devices. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services (MobiSys '20)*. Association for Computing Machinery, New York, NY, USA, 121–134. https://doi.org/10.1145/3386901.3388939

[33] Spotify. 2022. Storage information. https://support.spotify.com/us/article/storage-information/.

[34] Ke Sun, Ting Zhao, Wei Wang, and Lei Xie. 2018. VSkin: Sensing Touch Gestures on Surfaces of Mobile Devices Using Acoustic Signals. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. 591–605.

[35] Xiaoyuan Suo, Ying Zhu, and G Scott Owen. 2005. Graphical passwords: A survey. In *Proceedings of the 21st Annual Computer Security Applications Conference*. IEEE.

[36] Tuy Nguyen Tan and Hanho Lee. 2019. High-Secure Fingerprint Authentication System Using Ring-LWE Cryptography. *IEEE Access* 7 (2019), 23379–23387. https://doi.org/10.1109/ACCESS.2019.2899359

[37] Petroc Taylor. 2023. Global smartphone fingerprint sensor penetration rate 2014-2018. https://www.statista.com/statistics/804269/global-smartphone-fingerprint-sensor-penetration-rate/.

[38] Sebastian Uellenbeck, Markus Dürmuth, Christopher Wolf, and Thorsten Holz. 2013. Quantifying the security of graphical passwords: the case of android unlock patterns. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. 161–172.

[39] Cong Wu, Jing Chen, Kun He, Ziming Zhao, Ruiying Du, and Chen Zhang. 2022. EchoHand: High Accuracy and Presentation Attack Resistant Hand Authentication on Commodity Mobile Devices. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22)*. Association for Computing Machinery, New York, NY, USA, 2931–2945. https://doi.org/10.1145/3548606.3560553

[40] Cong Wu, Kun He, Jing Chen, Ziming Zhao, and Ruiying Du. 2020. Liveness is Not Enough: Enhancing Fingerprint Authentication with Behavioral Biometrics to Defeat Puppet Attacks. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, 2219–2236. https://www.usenix.org/conference/usenixsecurity20/presentation/wu

[41] Wencheng Yang, Song Wang, Jiankun Hu, Guanglou Zheng, and Craig Valli. 2019. Security and Accuracy of Fingerprint-Based Biometrics: A Review. *Symmetry* 11, 2 (Jan 2019), 141. https://doi.org/10.3390/sym11020141

[42] Yilin Yang, Yan Wang, Yingying Chen, and Chen Wang. 2020. EchoLock: Towards Low-Effort Mobile User Identification Leveraging Structure-Borne Echos. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (ASIA CCS '20)*. Association for Computing Machinery, New York, NY, USA, 772–783. https://doi.org/10.1145/3320269.3384741

[43] Yulong Zhang, Zhaonfeng Chen, Hui Xue, and Tao Wei. 2015. Fingerprints on mobile devices: Abusing and leaking. In *Black Hat Conference*.

[44] Yang Zhang, Peng Xia, Junzhou Luo, Zhen Ling, Benyuan Liu, and Xinwen Fu. 2012. Fingerprint Attack against Touch-Enabled Devices. In *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM '12)*. Association for Computing Machinery, New York, NY, USA, 57–68. https://doi.org/10.1145/2381934.2381947

[45] Bing Zhou, Jay Lohokare, Ruipeng Gao, and Fan Ye. 2018. EchoPrint: Two-factor Authentication using Acoustics and Vision on Smartphones. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. 321–336.