VibWrite: Towards Finger-input Authentication on Ubiquitous Surfaces via Physical Vibration

Jian Liu
WINLAB, Rutgers University
New Brunswick, NJ, USA 08901
jianliu@winlab.rutgers.edu

Chen Wang
WINLAB, Rutgers University
New Brunswick, NJ, USA 08901
chen.wang11@rutgers.edu

Yingying Chen*
WINLAB, Rutgers University
New Brunswick, NJ, USA 08901
yingche@scarletmail.rutgers.edu

Nitesh Saxena
University of Alabama at Birmingham
Birmingham, AL, USA 35294
saxena@uab.edu

ABSTRACT

The goal of this work is to enable user authentication via finger inputs on ubiquitous surfaces leveraging low-cost physical vibration. We propose *VibWrite* that extends finger-input authentication beyond touch screens to any solid surface for smart access systems (e.g., access to apartments, vehicles or smart appliances). It integrates passcode, behavioral and physiological characteristics, and surface dependency together to provide a low-cost, tangible and enhanced security solution. VibWrite builds upon a touch sensing technique with vibration signals that can operate on surfaces constructed from a broad range of materials. It is significantly different from traditional password-based approaches, which only authenticate the password itself rather than the legitimate user, and the behavioral biometrics-based solutions, which usually involve specific or expensive hardware (e.g., touch screen or fingerprint reader), incurring privacy concerns and suffering from smudge attacks. VibWrite is based on new algorithms to discriminate fine-grained finger inputs and supports three independent passcode secrets including PIN number, lock pattern, and simple gestures by extracting unique features in the frequency domain to capture both behavioral and physiological characteristics such as contacting area, touching force, and etc. VibWrite is implemented using a single pair of low-cost vibration motor and receiver that can be easily attached to any surface (e.g., a door panel, a desk or an appliance). Our extensive experiments demonstrate that VibWrite can authenticate users with high accuracy (e.g., over 95% within two trials), low false positive rate (e.g., less 3%) and is robust to various types of attacks.

CCS CONCEPTS

•Security and privacy → Authentication;

KEYWORDS

User authentication; finger-input; physical vibration; ubiquitous surfaces

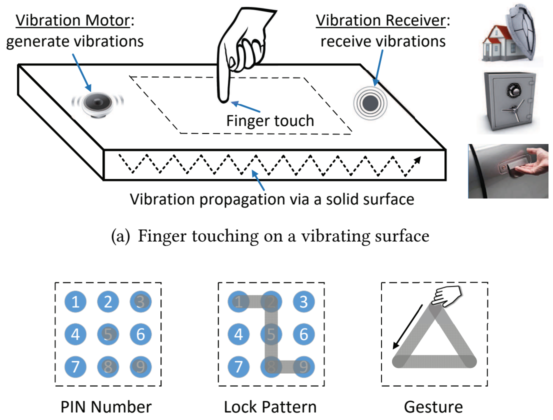
1 INTRODUCTION

The process of authentication verifies a user's identity and is frequently deployed at almost every corner of our daily lives. In particular, the increasingly wide deployment of *smart access systems*, which are defined as those used for keyless controlling access to corporate facilities/apartment buildings/hotel rooms/smart homes/vehicle doors, require the authentication process to play a broader role in numerous daily activities beyond the common form authentication on touch screen devices, such as mobile phones. The classic physical-key based access methods do not possess user authentication functionality. A market report shows that the deployment of smart security access systems is expected to grow rapidly at an annual rate of 7.49% and will reach a market value of $9.8 billion by the year of 2022 [1]. The current authentication process in smart security access systems mainly relies on traditional solutions supported by intercom, camera, card, or fingerprint based techniques. These approaches however involve expensive equipment, complex hardware installation, and diverse maintenance needs. The trend of employing low-cost low-power tangible user interfaces (TUI) to support user authentication in various facility entrances, apartment doors and vehicles has gained industry attentions recently. For example, token devices (e.g., smart ring, glove or pen) could be utilized for associating identities of their touch interactions [28, 46], and an ultra-thin sensing pad can be deployed in automobiles to perform driver authentication [7]. Moreover, isometric buttons appearing on new models of microwave ovens and stove tops and rotary inputs (e.g., used by iPod) can replace the regular physical buttons to provide better functionality and flexibility [2]. These new approaches appear promising of conducting user authentication and operating appliances/devices in smart systems leveraging capacitive sensing. However, these techniques require that the touched surface possesses electric conductivity and an electric field that produces/stores electrical energy, which largely limits the wide deployment of such solutions.

Along this direction, we start a new search in developing a low-cost general user authentication approach, which has the capability to work with any solid surface for smart access systems. The

[*] Corresponding Author.

CCS'17, Oct. 30–Nov. 3, 2017, Dallas, TX, USA.
© 2017 ACM. ISBN 978-1-4503-4946-8/17/10…$15.00
DOI: http://dx.doi.org/10.1145/3133956.3133964



(a) Finger touching on a vibrating surface

(b) Three types of secrets

Figure 1: Illustration of a finger touching on a solid surface under physical vibration, and three independent types of secrets for pervasive user authentication.

convenience of executing user authentication via touching any surface is enticing. For instance, a driver can just place his palm against the driver side window to access and start the vehicle. This has already been visualized in the popular movie ”Mission Impossible 5”, in which the featured BMW muscle car can be unlocked instantly when the lead actor pressed his palm against the side window. In another instance, a user can place his hand on the door panel of his apartment to perform authentication and unlock the entrance door without card access. Furthermore, electronic appliances in smart homes have a growing need to provide customized services for advanced safety needs such as prohibiting children and elderly people to operate risky appliances (e.g., oven and dryer), adjusting room temperature/lighting conditions and recommending TV content. A low-cost solution of tangible user authentication enabled on any solid surface could eliminate the need of installing touch screens on such electronic devices and make the customized services easy to deploy. Toward this end, our work seeks a general user authentication solution with smart access capability that can work with any solid surface (such as a door, a table or a vehicle's window), not limited to touch screens, and with minimum hardware and maintenance cost.

Existing Solutions. The traditional authentication solutions are based on passwords (i.e., texts and graphical patterns) [14, 16, 26, 40, 44]. However, all these approaches are based on the knowledge of the passwords, and thus suffer from password theft or shoulder surfing. Another direction of authentication involves physiological biometrics (e.g., fingerprints, iris patterns and face) [9, 18, 22, 24]. These mechanisms are less likely to suffer from identity theft. However, they usually require installation of expensive equipments and stir privacy concerns of the users. Furthermore, recent studies [15, 32, 36] allow users to rely on their familiar biometric-associated features (e.g., a sequence of 2D handwriting and corresponding pressure) extracted from mobile devices' sensitive touch screens instead of tedious passwords for user authentication. These approaches rely on touch screens, and are hard to be extended to general security access systems such as accessing corporate facilities, apartment buildings and smart homes when touch screens are not always available. In addition, oily residues, or smudges, on the touch screen surface may be used to recover user's graphical password (i.e., smudge attacks) [10].

Finger-input based Authentication over Any Surface through a Single Sensor. In this work, we introduce a new authentication system grounded on low-cost, low-power tangible user interface, called *VibWrite*, which has the flexibility to be deployed on ubiquitous surfaces. VibWrite leverages physical vibration to support authentication to emerging smart access security systems. To enable touching and writing on any surface during the authentication process, VibWrite builds upon a touch sensing technique using vibrations that is robust to environmental noise and can operate on surfaces constructed from a broad range of materials. As shown in Figure 1(a), when a vibration motor actively excites a surface resulting in the alteration of the shockwave propagation, the presence of the object or finger touching in contact with the surface can thus be sensed by analyzing the vibrations received by the sensor. VibWrite supports generalized vibration sensing based on a low-cost single sensor prototype that can be attached to solid surfaces (such as a door, a table or an appliance) and sense user touches and perform authentication flexibly from anywhere. By relying on the vibration signals in a relatively high frequency band (i.e., over $16kHz$), the system is hardly audible or distracting to the user, and is less susceptible to environmental interference from acoustic (i.e., mainly within a lower frequency band [41]) or radio-frequency noise. More importantly, vibration propagation is highly dependent on the surface material and shape in specific scenarios. VibWrite thus provides enhanced security by integrating location/surface uniqueness through such low-cost and tangible vibration-based user-interface. As another example, the vibration response of an office door is different from that of a house door. The unique behavioral information is embedded in both the behavioral biometrics as well as the surface being touched (e.g., the specific door in the office), making the system hard to be forged by attackers.

VibWrite provides users to choose from three different forms of secrets including PIN, lock pattern, and gesture (and signature in the future) to gain secure access as shown in Figure 1(b). The authentication process can be enabled on any solid surface beyond touch screens and without the constraint of the limited screen size. It is resilient to side-channel attacks when an adversary places a hidden vibration receiver on the authenticating surface or a nearby microphone to capture the received vibration signals. It is also robust to various adversarial activities, including the seemingly very powerful ones that observe the legitimate user's input multiple times and are aware of the passcode secret. It can authenticate the legitimate user and reject attacks well because of the following insights: 1) our study shows that vibration signals have the capability to perform cm-level location discrimination; and 2) unique features are embedded in a user's finger pressing at different locations on a solid surface. Such unique features reflect the characteristics of the user's finger touching on the medium (e.g., a door panel or a desk surface) including locations of touching, contacting area, touching force, and etc., making them capable to discriminate different touching locations of the same user and different users when touching on the same location. Thus, VibWrite enables users to finger-input (i.e., touch or write) on solid surface and is robust to passcode theft or passcode cracking by integrating 1) passcode, 2) behavioral and physiological characteristics (e.g., touching force and contacting area), and 3) surface dependency (e.g., house door or office desk) together to provide enhanced security. The main contributions of VibWrite are summarized as follows:

- We develop the first vibration-signal-based finger-input authentication system, which can be deployed on any solid surface for smart access systems (e.g., apartment entrances, car doors, electronic appliances and corporate desks).
- VibWrite captures intrinsic human physical characteristics presenting at specific location/surface for authentication through extracting unique features (e.g., frequency response and cepstral coefficient) in the frequency domain.
- VibWrite has the flexibility to support three types of secrets (i.e., PIN, lock pattern, and gesture) to meet different application requirements by developing new techniques of virtual grid point derivation, featured-based dynamic time warping (DTW) and distribution analysis based on earth mover's distance (EMD).
- VibWrite is implemented using a single pair of low-cost vibration motor and receiver, which involves minimum hardware installation and maintenance.
- We perform extensive experiments including authenticating legitimate users and modeling various types of attacks. The results demonstrate that VibWrite can effectively verify legitimate users with over 95% accuracy within two trials and less than 3% false positive rate.

2 RELATED WORK

User authentication becomes a critical step under the growing privacy concerns. Traditional user authentications utilize text-based passwords [26]. To ensure that a user's password cannot be easily guessed, the user has to memorize long strings of random characters, making it inconvenient [40]. Graphical passwords are proposed to ease the memory burden by letting users choose their pre-selected images from random choices of pictures [14, 16, 40] or Cued Clicked Points (CCP) in a sequence of images [12]. Additionally, grid lock pattern based approaches [25, 44] have been widely adopted to keep the user's mobile devices protected. Recent graphical authentication methods can resist shoulder surfing attacks by utilizing the Convex Hull Click Scheme [49] or the eye-gaze version of CCP [20]. However, these strategies eventually perform the authentication based on the knowledge of the passwords (e.g., text-based, image-based and lock pattern-based) and cannot tell whether the password is entered by the legitimate user or not.

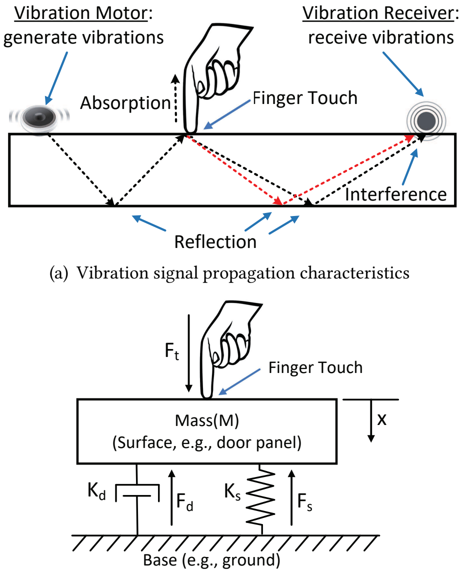
To ensure that the secret inputs used for authentication are physically from the legitimate user, biometrics-based schemes (e.g., fingerprints [9], iris patterns [22], retina patterns [24], and face [18]) have been drawn considerable attention recently. However, physiological biometrics are sensitive personal information, which may involve privacy concerns, thus are not widely accepted. To reduce the privacy concerns, a compromised approach is to authenticate users based on their behavioral characteristics, including unique keystroke dynamic [33], mouse movements [50], and gait patterns [31]. Although these approaches are less sensitive in terms of privacy, they are designed for continuous user verification during the period that the user operates the keyboard, moves a mouse or takes a walk, rather than one-time authentication.

To provide authentication to the emerging smart access systems needed by corporate facilities, apartment buildings, hotel rooms, and smart homes, techniques involving intercom [29], camera [43], access card [34] and fingerprint [9] have been explored. For example, KinWrite [43] uses Kinect, a vision-based platform, to capture the user's 3D handwriting patterns for authentication. These approaches usually involve expensive hardware, complex installation process, and diverse maintenance efforts. Recent studies successfully combine 2D handwriting and behavior features such as corresponding writing pressure, writing speed, and correlation between multiple fingers on touch screens to provide enhanced security [15, 32, 36]. The limitation is that the authentication relies on touch screens, which may suffer from smudge attacks [10] and are not always available in smart access systems. Toward this end, we propose VibWrite that extends the authentication process beyond touch screens to any solid surface leveraging vibration signals. VibWrite will have the authentication capability in a broad array of applications including entry access (e.g., smart building, car doors) and supporting customized services in appliances and devices at smart homes. The authentication process combines password and human physical traits, and supports three types of secret independently including PIN, lock pattern, and gesture input for emerging smart access systems.

3 PHYSICAL VIBRATION PROPAGATION

Physical vibration is a mechanical phenomenon, which creates a mechanical wave transferring the initial energy through a medium. Similar to the transmission of wireless signals, when a vibration signal travels through a medium, it experiences attenuation along the propagation path and reflection/diffraction when the signal hits the boundary of two different media (e.g., the contacting area between a finger and a medium). Figure 2(a) illustrates the reflection and diffraction of a vibration signal propagating in a solid surface when a finger touches the area in between the vibration signal generator and receiver. When the vibration signal hits the contacting area of the finger, part of the signal reflects back to the surface and the rest of it propagates into the finger (i.e., absorption) and bounces back to the surface along a different propagation path. The vibration signal is affected by the touching location of the finger and traverses different paths before reaching the receiver (i.e., vibration sensor). Thus, the touching location information is embedded in the various interference effects captured at the receiver.

Furthermore, when a finger touches the surface of an object (e.g., a table), the flexibility of the object is affected not only by the touching location but also the strength of touch. A recent study [45] utilizes these properties to enable a commodity phone to recognize the force applied to its phone body and screen. To mathematically model the vibration effect on the object under an external force caused by the finger touch, we consider a spring-mass-damper system as shown in Figure 2(b). A free body diagram with the mass M represents the vibrating surface, while the external force F_t is caused by the finger touch. Moreover, the vertical shaft has an



(a) Vibration signal propagation characteristics

(b) Vibration model under finger touch

Figure 2: Illustration of the propagation characteristics of vibration signals on a solid surface.

effective spring constant of K_s and a damping coefficient of K_d. When the surface has a vertical displacement of x, we have

$$F_t = K_d(\frac{d}{dt})x + K_s x + M(\frac{d}{dt})^2 x. \tag{1}$$

To satisfy the equilibrium condition, the vertical displacement x is dependent on the external force F_t. This indicates that the finger touching force could be captured by analyzing the received vibration signals and utilized as a biometric-associated feature in VibWrite. Note that the above analysis also works on vertical planar surface (e.g., door panel) as the equilibrium condition could be analyzed along the direction perpendicular to the surface.

In addition, Dong *et al.* [17] experimentally demonstrate that the vibration energy absorbed into the human finger-hand-arm system is different under different vibration frequencies. In our empirical study we find that the frequency response of the same user finger-press presents higher correlation than that of different users when they touch the same location on a surface. This important observation suggests that the vibration propagation properties are strongly influenced by unique human physical traits such as contacting area, touching force and etc., which can assist ubiquitous user authentication together with passcode on any surface beyond touch screens.

4 APPROACH OVERVIEW

In this section, we present the attack model and system overview of VibWrite.

4.1 Attack Model

We consider the following attacks that are harmful to the proposed ubiquitous authentication functionalities.

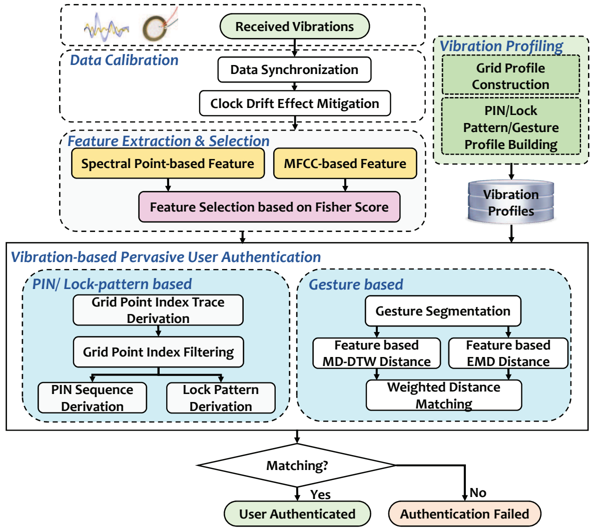

Received Vibrations
Data Calibration
Data Synchronization
Clock Drift Effect Mitigation
Feature Extraction & Selection
Spectral Point-based Feature
MFCC-based Feature
Feature Selection based on Fisher Score
Vibration Profiling
Grid Profile Construction
PIN/Lock Pattern/Gesture Profile Building
Vibration Profiles
Vibration-based Pervasive User Authentication
PIN/ Lock-pattern based
Grid Point Index Trace Derivation
Grid Point Index Filtering
PIN Sequence Derivation
Lock Pattern Derivation
Gesture based
Gesture Segmentation
Feature based MD-DTW Distance
Feature based EMD Distance
Weighted Distance Matching
Matching?
Yes
No
User Authenticated
Authentication Failed

Figure 3: Overview of VibWrite architecture.

Blind Attack. An adversary randomly touches on the authentication surface equipped with the VibWrite system, hoping the random touching events can result in similar impacts to the vibration signals as the legitimate user does and passes the authentication.

Credential-aware Attack. An adversary has the prior knowledge of the legitimate user's credentials, including the PIN number, lock pattern or personal gesture, but does not possess the knowledge of the VibWrite setting details such as the grid size, gesture region, and the authentication surface.

Knowledgeable Observer Attack. An adversary is capable of both observing the legitimate user's hand movements when he is passing the authentication system via shoulder surfing or video taping as well as knowing the user's credentials and VibWrite setting details. The adversary tries to imitate the legitimate user's hand or finger movements based on his understanding of the user's credentials to pass the authentication.

Side-channel Attack. An adversary makes an effort to hack the VibWrite system directly in the hope of capturing the similar vibration signals of the legitimate user by placing a hidden vibration receiver on the authentication surface or employing a microphone in a nearby location.

4.2 System Overview

The basic idea underlying VibWrite is to analyze unique features from the received vibration signals to enable authentication on ubiquitous object surfaces such as entrances (e.g., apartment building or car doors) and smart home appliances (e.g., hot stove and dryer). In particular, VibWrite can be triggered when a person moves closer to the security access area (e.g., a door panel), which can be easily achieved using low power proximity sensors or motion sensors [37, 39]. As illustrated in Figure 3, the vibration motor then generates low annoyance vibrations and VibWrite starts taking inputs of vibrational signals from the vibration receiver. The system first performs *Data Calibration* (Section 5.2) including data synchronization and clock drift effect mitigation to ensure the received vibration signals always synchronized and eliminate the effects caused by the clock drift (i.e., inconsistent sampling frequency).

VibWrite then extracts and selects vibration features (Section 5) in the frequency domain from the synchronized vibration signals within a sliding window. We find that *Spectral Point-based Feature* (i.e., frequency amplitude of each spectral point) and *MFCC-based Feature* (Mel-frequency cepstral coefficient [27]) reflect the intrinsic physical traits embedded in the user's finger inputs. The system further performs *feature selection* based on the Fisher Score [19] on top of the Spectral Point-based and MFCC-based features by selecting a subset of features exhibiting more discriminative power among different touching locations as well as maintaining feature consistency within each touching location.

The extracted vibration features are used by two phases in VibWrite: *profiling* and *authentication*. In both PIN number based and lock pattern based authentications, a grid is drawn on the touching surface. In the profiling phase, the features are extracted and captured while a user first enrolls in the system and presses his finger at different grid points on the touching surface. These features are labeled and saved to build the user's profile in *Grid Profile Construction*.

During the authentication phase, the received vibration signals are utilized to extract vibration features. The extracted features then serve as inputs to *Grid Point Index Trace Derivation* via a classifier based on Supporting Vector Machine (SVM) trained by the grid profiles. The classifier compares the extracted features with the stored ones in the profile to filter out the signal segments before and after the finger inputs and derive grid point trace containing finger touching inputs. The derived grid point trace would then be put into *Grid Point Index Filtering* (Section 6.2) to eliminate the incorrectly classified grid point indices and obtain the ones corresponding to the finger presses in the grid point index trace. Next, the filtered grid point trace would be recovered to the PIN sequence/lock pattern via *PIN Sequence Derivation* or *Lock Pattern Derivation* (Section 6.3). The recovered PIN number/lock pattern is then compared with the local stored PIN/lock pattern information for the final authentication.

Independently, VibWrite also enables the user to perform simple gestures (e.g., drawing a circle on the surface) for authentication without the restrictions of pressing/passing the grid points on the authentication surface. Different from the fixed grids in PIN/lock pattern based authentication, using gestures provides more flexibility for authentication. However, even for the same user, the same finger gesture could be slightly different at different authentication times due to the lack of consistency. Thus, the mechanism for gesture-based authentication in VibWrite needs to capture the intrinsic gesture behavior to deal with gesture inconsistency while preserving individual diversity. In particular, during the gesture-based authentication, VibWrite first identifies the signal segment containing the gesture operation via *Gesture Segmentation*. In the profiling phase, the extracted feature sequence (i.e., Spectral Point-based and MFCC-based features) from the gesture segments are saved to build the specific user's profile. To measure the similarity of generated features in the authentication phase to the gesture profiles, VibWrite addresses the gesture inconsistency problem by considering both time warped feature sequences and the distribution of the features. This is achieved by calculating both MD-DTW (Multi-Dimensional Dynamic Time Warping) Distance [42] and EMD (Earth Mover Distance) [35] of the extracted feature sequences to the profiles. The weighted distance combination in *Weighted Distance Matching* obtains the combined distance from the two techniques. Finally, VibWrite makes decision as user authenticated or access denied by checking a threshold to the calculated distances between input gestures and the stored profiles.

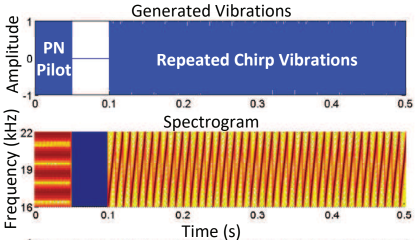


Figure 4: Example of generated vibrations between $16kHz$ **and** $22kHz$**.**

5 VIBRATION SIGNAL DESIGN AND FEATURE EXTRACTION & SELECTION

In this section, we first describe the details of vibration signal design and calibration. We then present how to extract and select unique features for the authentication process in VibWrite.

5.1 Vibration Signal Design

To facilitate finger-input based user authentication via physical vibration, the vibration signals used in our system need to contain a broad range of frequencies to increase the diversity of vibration features in the frequency domain. Specifically, we generate repeated chirp vibration signals to linearly sweep frequency from $16kHz$ to $22kHz$, which are hardly audible to most human ears [21]. Additionally, such frequency range is much higher than the frequency range of ambient noise and the vibrations caused by human body (e.g., breathing and heart beating). This makes our system less possible to be interfered by these unrelated noises. Figure 4 illustrates an example of the generated vibration signal and its corresponding spectrogram.In particular, there is a short pseudo-noise (PN) sequence preamble played before the repeated chirp vibrations, which is used for the signal synchronization. We leave the details in Section 5.2. After transmitting PN pilot, with a $50ms$ pause, the vibration motor repeatedly transmits the chirp vibration signal to keep its continuous sensing capability while performing authentication. The length of each chirp vibration signal is set to T=$10ms$, which provides high time resolution to enable continuously finger-input sensing.

5.2 Vibration Signal Calibration

Vibration Signal Synchronization. The timing of the VibWrite's vibration motor and receiver needs to be synchronized, so that we could guarantee that each sliding window being used to extract vibration features contains the same parts of the chirp vibration signals without time delay. Therefore, they can be used for further comparison of their extracted features and capture the difference

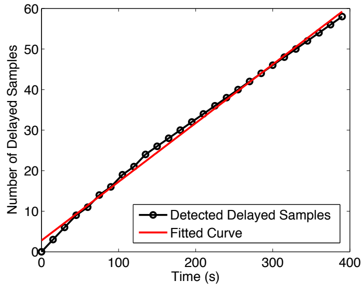


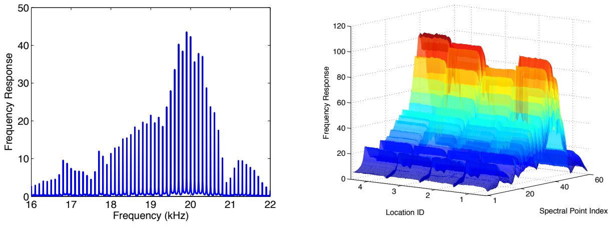
Figure 5: Illustration of clock drift effect mitigation.

in each window when the finger touches different positions on the surface. In order to avoid the uncertainty, we add a pseudo-noise (PN) sequence preamble (i.e., 2400 samples) [38], which has ideal autocorrelation properties, at the beginning of the generated chirp vibration signals as illustrated in Figure 4. We then synchronize the received vibrations using cross-correlation between the PN sequence of the received vibration signal and the known generated PN sequence.

Clock Drift Effect Mitigation. When the vibration receiver senses the vibration, the analog voltage signals created by the sensor will be converted into the digitized signals via an Analog to Digital Converter (ADC). The ADC can be configured at a wide range of rates, and it is usually set to sample the analog signals at a fixed frequency driven by different application requirements. For instance, a few options (e.g., $32kHz$, $44.1kHz$ and $48kHz$) can be set in most smartphones' audio ADCs in terms of the required audio recording quality. However, we experimentally find that the sampling rate may be not a fixed value over time due to imperfect clock, and there exists a small gap between the real sampling rate and the configured sampling rate. To eliminate the effect caused by the clock drift, we estimate the sampling rate offset during a short calibration phase at the beginning. During the calibration, the vibration motor periodically sends a short vibration chirp with a fixed time interval (e.g., $2s$). The time intervals between these chirps should be fixed value as well if there is no clock drift. We use cross-correlation to measure the sample delays of the received vibration chirps over time, which is illustrated in Figure 5. We observe that the number of the delayed samples increases linearly over time, indicating that the real sampling rate is slightly larger than the configured sampling rate but remains a relative fixed value. We then use a least-squares based approach to fit a quadratic curve to the measured delayed samples, and obtain the slop k to shift the starting point S_p of each received vibration chirp to $S_p = S_p - \lfloor kt \rfloor$, where t is the time interval between the current vibration chirp and the first received vibration chirp.

5.3 Spectral Point-based Feature Extraction

In order to extract unique vibration features from the received vibrations to discriminate the finger touches on different surface locations and distinguish different users touching a same surface location, we first analyze the received vibration signals in the frequency domain using a $200ms$ sliding window. Figure 6(a) presents an example of the Fast Fourier Transform (FFT) of a time series of the received vibration signals, ranging from $16kHz$ to $22kHz$, in a sliding window. The transmitted chirp vibration signal has fundamental frequencies that are all multiples of the frequency $1/T$ Hz, where T is the time duration of each chirp vibration signal (e.g., $T = 0.01s$ in VibWrite). We find that the amplitudes of some designated frequency components in the signals (i.e., peak values in Figure 6(a)), called *spectral points*, are most sensitive to the minute changes caused by finger touching or swiping. These spectral points are more sensitive to the finger touches and could be utilized to differentiate different surface locations finger presses or finger moving along. For example, in our preliminary experiments, the vibration signals are collected when a user's finger presses at four different locations of a solid surface (i.e., wooden table) equipped with our vibration motor and receiver. We observe obvious distinguishable patterns of the frequency amplitude at these 60 spectral points (i.e., $\frac{22000-16000}{100} = 60$) between different locations, which are shown in Figure 6(b). Furthermore, the spectral points in the frequency domain may not be exactly spaced at $100Hz$ due to imperfect sampling module. We thus design a threshold-based strategy (i.e., minimum distance between two neighboring peaks and minimum height of each detected peak) to find peaks of the frequency response to extract each spectral point feature.



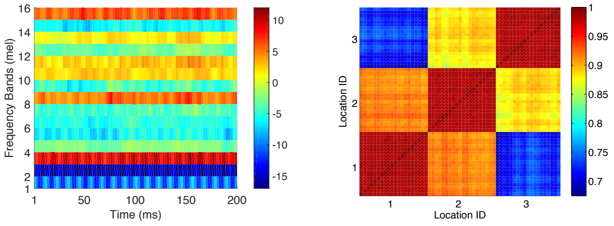
(a) Spectral points at every $100Hz$ interval

(b) Distinguishable spectral points when a finger presses 4 different locations

Figure 6: Illustration of the frequency response of the received vibrations in a $0.2s$ time window. And the frequency response is depicted at spectral points when a finger presses 4 different locations of a desk.

5.4 MFCC-based Feature Extraction

The Mel-frequency cepstral coefficient (MFCC) is widely used to represent the short-term power spectrum of acoustic or vibration signals [27] and can represent the dynamic features of the signals with both linear and nonlinear properties. While the MFCCs are able to distinguish people's sound differences in speech and voice recognition, we find that they can also characterize the vibration signals transmitting via the medium of a solid surface on which the user's finger touches, because the user's behavioral and physiological characteristics (e.g. touch area and pressure) and the touching position can cause different changes to the vibration propagation. We thus extract the MFCC-based features to characterize the different vibration signatures when the user touches or writes at different positions on the surface. In particular, we calculate the MFCCs of the received vibration signals in each sliding window. The number of filterbank channels is set to 32, and 16-th order cepstral coefficients are computed in each $20ms$ Hanning window, shifting $2ms$ each time.



(a)

(b)

Figure 7: MFCC feature illustration: (a) Example of the extracted MFCC features and (b) Pearson Correlation between MFCC features when a finger presses three different locations on a desk surface.

Figure 7(a) shows the MFCCs extracted from the received vibration signals in a $0.2s$ sliding window when the user presses on a solid surface. We observe that the extracted MFCCs have a periodical pattern, which is caused by the cycle of the repeated vibration chirp signals. Figure 7(b) shows Pearson correlation coefficient [8] of the MFCC-based features when the user's finger touches at three different locations. In this experiment, twenty consecutive sliding time windows (i.e., instances) are used to extract MFCCs for each finger-touching location to compare the similarity between different finger touches. We observe that the MFCC features of the same finger-touching location present higher correlation than that of different locations, which confirms the effectiveness of utilizing the MFCC features to characterize the user's finger-touching on the surface.

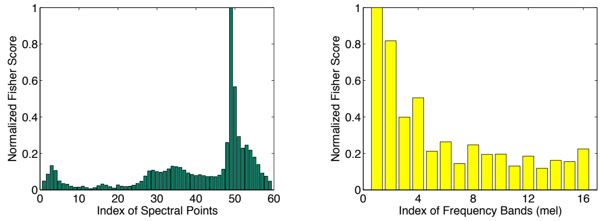
5.5 Feature Selection based on Fisher Score

From our experiments, we observe that not all extracted features including both spectral points and MFCC are unique enough to discriminate different touching locations and distinguish different users touching the same location. The discrimination power is dependent on the extracted features at specific frequencies or Mel-frequency bands. We therefore propose to select features based on Fisher Score [19] to find a subset of features which are more distinct between classes (i.e., touching locations per user) and consistent within a class. The fisher score of the r-th feature candidate is defined as follows:

$$F_r = \frac{\sum_{i=1}^{c} n_i(\mu_i - \mu)^2}{\sum_{i=1}^{c} n_i \delta_i^2}, \qquad (2)$$

where n_i is the number of instances in class i. And μ_i and δ_i^2 denote the mean and variance of class i, $i = 1, \ldots, c$, corresponding to the r-th feature candidate. μ denotes the mean of r-th feature candidates in the whole data sets.

To analyze the feature difference between different frequency bands, we consider each spectral point or MFCCs at each frequency band as an individual feature candidate. Figure 8 shows the normalized fisher scores of both the spectral point based and MFCC based features that we use to perform user authentication. In VibWrite, we empirically choose top 30 spectral point based features, and top 8 MFCC based features which are more sensitive to the finger pressing and swiping.



(a) Fisher score of spectral point based features

(b) Fisher score of MFCC based features

Figure 8: Fisher score of the feature candidates (a) spectral point based and (b) MFCC based.

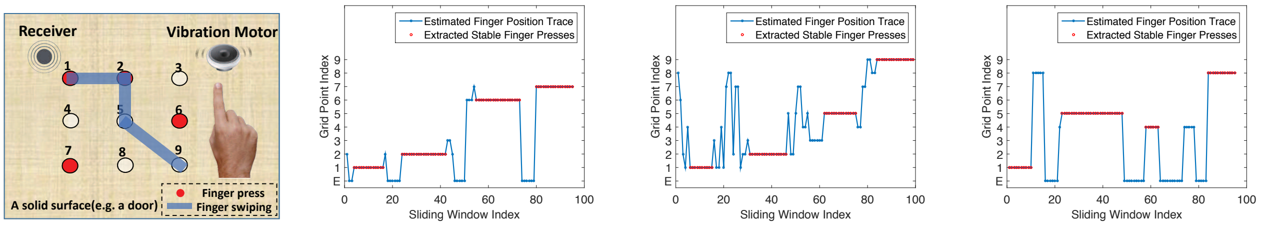
6 AUTHENTICATION USING PIN NUMBERS AND LOCK PATTERNS

The VibWrite system allows users to perform PIN number based authentication by touching grid points on a solid surface or conduct lock pattern based authentication by swiping finger through the grid points. Depending on the type of applications, the solid surface could be a range of options including an apartment door, a car door, an executive's office desk or a smart appliance. VibWrite first converts the received vibration signals to a time series of grid point indices, then filters out the incorrectly classified grid point indices and finally determines the PIN sequence/lock pattern based on the derived grid point indices.

6.1 Deriving Grid Point Index Traces

The system takes the received vibration signals as input when the user enters PIN sequence/lock pattern. In particular, we apply a sliding window to the vibration signals and derive vibration features (e.g. spectrum-based feature and MFCC-based feature) in every sliding window. We then apply a machine learning-based grid point classifier based on the Support Vector Machine (SVM) using LIBSVM [11] to estimate the finger-press positions in terms of the grid point index for each sliding window, by leveraging the user's personal grid profile. The resulted grid point index trace is actually an estimated finger-press position trace which reflects the finger position changes among the grid point indices in the entire PIN sequence/lock pattern input duration. Note that when we derive grid point index trace, it involves user's behavior and physical characteristics. It is highly difficult for an unauthorized user to obtain correct grid point index at this step because the system needs to compare with the authorized user's profile, which integrates both PIN/Lock pattern and the user's behavior characteristics. Based on the derived grid point index trace, we can recognize the user's PIN sequence/lock pattern input and verify their identities.

Figure 9 shows an example of the user's PIN sequence/lock pattern based authentication on a solid surface (e.g. an apartment door) with a 3×3 grid. The predesigned grid is drawn in-between the receiver and vibration motor as shown in Figure 9(a), and the distance between the grid points is $3cm$. The user first builds a personal grid profile, which is discussed in Section 6.4. The user then presses the grid points "1267" sequentially to input a PIN sequence and swipes the finger through the grid points "1-2-5-9" to input a lock pattern as shown in Figure 9(a). The vibration features during the PIN sequence/ lock pattern input are extracted in each



(a) User presses a PIN sequence "1267" and swipes a lock pattern "1 − 2 − 5 − 9" on a 3 × 3 grid

(b) Estimated finger position trace in terms of grid point index when the user enters the PIN sequence "1267"

(c) Estimated finger position trace in terms of grid point index when the user swipes the lock pattern "1 − 2 − 5 − 9"

(d) Example of an attacker entering the legitimate user's PIN sequence "1267" on the same grid of the same desk surface.

Figure 9: Example of PIN sequence/lock pattern derivation in sliding windows when entering a PIN sequence/lock pattern on a solid surface.

sliding window and are inputted to the SVM-based classifier. The estimated finger position trace (i.e., grid point index trace) for the PIN sequence input "1267" is shown in Figure 9(b). We observe that when the user presses on a number with the finger staying on the virtual key, the consecutive same grid points corresponding to the key can be obtained, and when the user moves the finger in the air to the next key, the vibration signals are classified as "E" representing "Empty" based on the vibration profile collected when no finger presses the surface.

Figure 9(c) shows the estimated finger position trace of the lock pattern "1-2-5-9". We observe that when the finger swipes near a virtual key, the vibration signals will be classified to the corresponding grid point index. In particular, the consecutive same grid points can be obtained for the duration beginning from the finger moving close to, pressing on, to just swiping away from the virtual key. Thus the derived grid point index trace can reflect the user's finger positions on the grid and can be utilized to further derive the user's PIN sequence/lock pattern inputs.

6.2 Grid Point Index Filtering

However, the derived grid point index traces contain incorrectly classified grid point indices, which are due to the unstable vibration features caused by the varying finger touching area and force when the finger is just detaching or pressing on the surface (e.g., the noises in Figure 9(b)), or are because the swiping finger is far from any of the predesigned profiled virtual keys (e.g., the noisy indices in Figure 9(c)). These incorrectly classified grid point indices should be excluded when deriving the passcode patterns.

We develop a grid point index filter to determine the segments that have consecutive same grid point indices. Intuitively, these segments are corresponding to the time periods when the user's finger is pressing on or swiping near a grid point, which means they are more reliable results for identifying the PIN sequence/lock pattern. The grid point index filter consists of three steps: 1) calculating the difference between every two consecutive grid point indices in the trace and the firm presses will generate consecutive "0" for the differential grid point index; 2) searching for the starting and ending points of the consecutive differential grid point indices (i.e., 0s) to extract *finger-press segment*, indicating the finger positions of the firm finger presses right on or near virtual keys; 3) removing the grid point indices from the trace that are out of the finger-press segments. The red dots in Figure 9(b) and Figure 9(c) are filtered grid point indices for the PIN sequence and lock pattern derivation, respectively.

6.3 PIN Sequence/Lock-pattern Derivation

Next, we further confirm each finger-press segment based on their time length and remove the incorrect finger location estimations to derive the PIN sequence/lock pattern. The intuition is that when users enter their PIN sequences, the finger press for each PIN number lasts for a certain amount of time. And when users draw their lock patterns, the duration beginning from the finger swiping close, right pressing on, to finger swiping away from each virtual key should last for an amount of time. The grid point index segments shorter than this amount of time are highly possible to be incorrect finger location estimations. We empirically determine the threshold of minimum finger-press duration (i.e., 300ms) to remove the finger-press segments with shorter time duration. Finally, given the length of the user's PIN sequence/lock pattern, the system finds the same number of the longest finger-press segments as the valid finger-press segments and derives the PIN sequence/lock pattern by mapping the segments' grid point indices to the virtual keys.

6.4 Grid Profile Construction

We notice that the users can generate individually unique vibration features even by pressing at the same position of a solid surface due to the individual's different behavioral and physiological characteristics (i.e., touching area and pressure on the surface). The user's such unique vibration features can provide another level of security to our user authentication in addition to the secrecy of passcodes.Our PIN/Lock-pattern based authentication requires constructing the user's profile corresponding to every grid point, which enables successful identification of the input virtual keys during authentication. Specifically, the VibWrite system records a short time period (e.g., 1 to 5 seconds per grid point) of received vibration signals when the user presses at each grid point. The recorded vibration signals are used to derive the vibration features in sliding windows. The feature in each sliding window is labeled with corresponding grid point index. In addition, we also build a profile when no finger touches the surface and label it as "E" (i.e.,"empty") to discriminate whether finger presses on the surface.

To illustrate the security provided by the user's unique vibration features in addition to the passcodes for PIN number/lock pattern based authentication. We ask an attacker to enter the legitimate

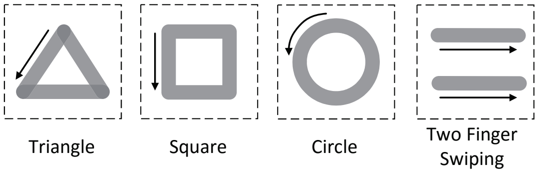


Figure 10: Illustration of the four pre-defined finger gestures for gesture-based authentication.

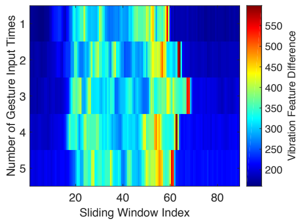


Figure 11: Illustration of gesture segmentation when a user inputs gestures for five times.

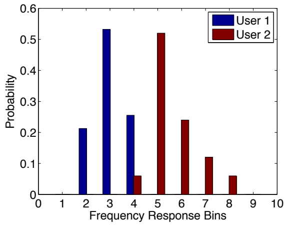


Figure 12: Histogram of frequency response at a spectral point for two users swiping a same gesture.

user's same PIN number "1267" via VibWrite on the same grid and the same surface as shown in Figure 9(a). The VibWrite processes the attacker's vibration signals based on the legitimate user's grid profile and the results are shown in Figure 9(d). We observe that nearly all the vibration features of the attacker are incorrectly classified and thus cannot pass the authentication, which verifies the effectiveness of the individual physical characteristics contained in the user's grid profile.

7 AUTHENTICATION USING GESTURES

Different from PIN/lock pattern based authentications, using gestures provides more flexibility for authentication. In particular, VibWrite defines four simple finger gestures as shown in Figure 10: swiping a single finger along three patterns including a triangle, square and circle, and swiping two fingers horizontally.

7.1 Gesture Segmentation

To facilitate the gesture-based authentication, our system needs to first detect the occurrence of the user's gesture input from the received vibration signals and remove the vibration signals with no gestures (i.e., no touch on the surface). Specifically, VibWrite extracts vibration features from spectral points and MFCC and then calculates vibration feature differences between the received vibration signals and those in the profile when no finger touches on the surface. The intuition is that when the user inputs a gesture, the finger swipes on the surface, causing the vibration features to differ largely from those when there is no finger touching. Figure 11 shows an example of calculated vibration feature differences when the user inputs square gestures on the surface for five times. For all the five gesture inputs, we observe the vibration feature difference grows higher (e.g. over 300) when the finger swipes on the surface and falls back to lower values (e.g., around 200) when the finger releases from the surface. We thus normalize the vibration feature differences and segment each gesture via a threshold.

7.2 Distance Calculation of Feature Sequence

User authentication using such simple gestures is much harder due to lack of unique secrecy to discriminate different users. Moreover, the speed, duration, and trajectory of the same user's gestures could be different from time to time, which causes gesture inconsistency and makes the generated vibration signals present different lengths and results in varying density of locations within the swiped pattern. In addition to feature extraction containing user's unique physical traits, we resort to two techniques to complete the authentication process in high accuracy to cope with these challenges: the Dynamic Time Warping (DTW) [42] is exploited to deal with gesture inconsistency, and the earth mover's distance (EMD) [35] technique is employed to preserve individual diversity because the feature distribution of the same user should have a higher similarity than that from different users.

Specifically, we first derive a time series of vibration features based on the vibration signals in segmented gestures using a sliding window. The DTW technique stretches and compresses required parts to allow a proper comparison between two data sequences. Therefore, it is useful to compare the vibration feature traces extracted from two segmented gestures regardless of different swiping speeds. In our system, vibration features are in a format that reports both frequency amplitude at multiple spectral points and MFCC coefficients, which is discussed in Section 5. To perform multidimensional sequence alignment, our system applies Multi-Dimensional Dynamic Time Warping (MD-DTW) [42], in which the vector norm is utilized to calculate the distance matrix according to:

$$d(v_i, v'_j) = \sum_{p=1}^{P} (v_i(p) - v'_j(p))^2, \tag{3}$$

where $V = v_1, v_2, ..., v_T$ and V'=$v'_1, v'_2, ..., v'_T$ are two vibration feature traces for gesture discrimination, and P is the number of dimensions of the sequence data (i.e., the number of extracted features within each window). A least cost path is found through this matrix and the MD-DTW distance is the sum of the matrix elements along the path.

Besides time warped feature sequence, we find that the histogram of the spectral point based features preserve individual diversity and can be used to distinguish different users when even the same gesture is swiped. Figure 12 shows the feasibility study results where two users swipe their fingers following an exactly same circle gesture pattern on a desk surface. The histogram of frequency response (quantized to 10 bins) at a specific spectral point during their swiping presents distinct distributions that can clearly distinguish these two users. We thus take the advantage of the EMD-based distribution difference to preserve the individual diversity during gesture based authentication. Specifically, we normalize the EMD distance and MD-DTW distance to be integrated for final authentication. If the integrated distance to the gesture profiles is larger than a threshold, VibWrite regards the swiped gesture as an unknown gesture and fails the authentication. Otherwise, we consider the swiped gesture is from the user whose profile results in the minimum integrated MD-DTW and EMD distance.

7.3 Gesture Profile Construction

Unlike grid point profile construction, VibWrite does not need to construct profiles for each grid point for the gesture-based authentication. Instead, when constructing the gesture profile for a particular user, VibWrite collects the vibration signals while the user swipes a finger following a predefined gesture. In particular, we use the sequence of the vibration features extracted from the segmented signals for building individual gesture profile. Though the profile only contains simple gestures, such profile contains the user's unique behavior and physiological characteristics and is sufficient to perform user authentication. We also build a profile with the vibration signals when there is no finger touching on the surface to determine the presence of finger touching or not for gesture segmentation.

8 PERFORMANCE EVALUATION

In this section, we first describe the experimental setup and methodology. We then present the performance of VibWrite in terms of authenticating the legitimate user and its robustness under various attacking scenarios.

8.1 Prototyping and Experimental Setup

We evaluate the performance of user authentication using PIN and lock patterns on a 3×3 square-shaped grid. In practice, the grid patterns could be flexibly extended as needed. The grid is drawn on a solid surface in a typical office environment. The distance between every two adjacent grid points is $3cm$. We test with two different surfaces as shown in Figure 13: one with the testing region resided below the vibration motor and receiver on a wooden table (e.g., the executive's desk in a company), and the other with the testing region resided in between the motor and receiver on a door panel (e.g., an apartment door). For the user authentication using gestures, we remove the restriction of pressing/passing the grid points on the authentication surface, and aim to utilize the simplest finger gestures as shown in Figure 10. We want to demonstrate that even the simplest finger gestures carry the unique behavioral and physiological characteristics reflected by the physical vibrations. The gesture patterns are drawn on the table within a $6cm \times 6cm$ region between the vibration motor and receiver to guide user's swiping.

The vibration generator is implemented with a Linear Resonant Actuator (LRA) based motor, which has a wide frequency response. The frequency and amplitude of the generated vibration can be regulated by the frequency and peak-to-peak voltage of an input analog signal. The low-cost vibration receiver is implemented with a vibration receiver (i.e., piezoelectric sensor) and a low-power consumption amplifier, which can be easily plugged into the standard audio jack of any audio recording device (e.g., mobile phone) to sense vibration signals. The sampling rate of the vibration receiver is determined by the audio recording device, which is typically $48kHz$. The size of vibration motor and receiver is very small, which makes them easily to be attached to any solid surface. Compared to other authentication systems based on cameras, touch screens, or biometric readers, in VibWrite we seek to explore using low-cost sensor settings (i.e., vibration motor and receiver) for the potential of wide-deployment such as in apartment buildings, hotel rooms, smart homes, office desks, etc. Besides the vibration motor and receiver, our system needs additional supporting hardware including, but not limited to, amplifier, ADC, micro-controller and storage device to perform necessary data process, feature extraction and profile matching. With these required components, we roughly estimate the cost of an end-to-end system could be maintained around tens of dollars (e.g., \$20 ~ \$50). As a comparison, some existing authentication systems (e.g., face recognition based and fingerprint based [3, 5, 6]) may usually cost hundreds of dollars.

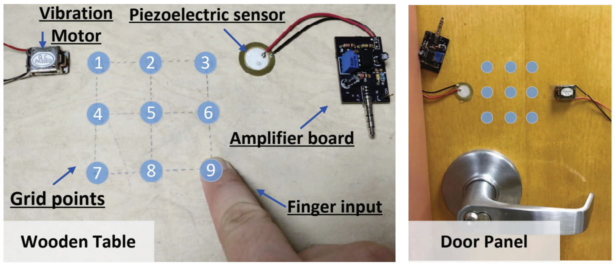


Figure 13: Experimental setup of VibWrite on a wooden table and door panel.

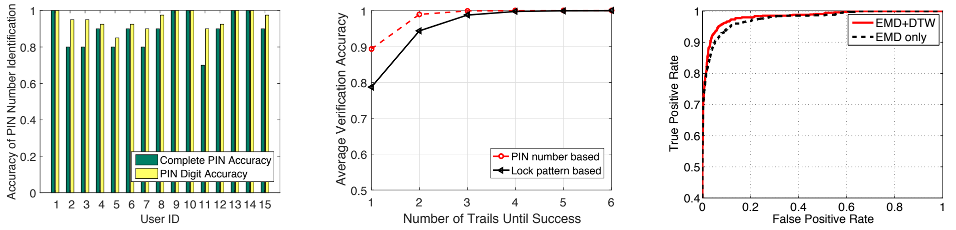
8.2 Evaluation Scenarios & Data Collection

8.2.1 Legitimate User Verification. We recruit 15 participants to evaluate the performance of VibWrite under three types of authentication. [1] Our data is collected across three-month period, and 15 participants were involved across different days. Additionally, before the data collection, we allow users to practice multiple rounds of authentication inputs on the authenticating surface to get familiar with the VibWrite system. 1) For PIN number based authentication, each user is asked to sequentially press the 9 grid points for $5s$ to create his/her grid profiles. During verification, each user presses 10 random 4-digit PIN sequences as their passcodes. 2) For lock pattern based authentication, our system uses the same grid point profiles. During testing, each user swipes his/her finger through 10 lock patterns to verify the system's authentication performance. 3) For gesture based authentication, each user chooses one of the four gestures as shown in Figure 10 as their preferred gestures and swipes the finger gesture 10 times. In total, we collected 450 genuine input passcodes (i.e., PIN sequences, lock patterns and gestures) for each motor/receiver placement to evaluate legitimate user access authentication. We further collected attack data to evaluate the VibWrite performance under attack scenarios.

8.2.2 Various Attack Scenarios. We evaluate the robustness of VibWrite under various types of attack. Specifically, we choose one user as a legitimate user and the rest users as attackers to launch the attacks.

Blind Attack. The attacker randomly guesses the legitimate user's PIN, lock pattern and gesture and uses his/her finger to press and swipe on the solid surface for 10 times. In total, we collected 420 blind attack inputs.

[1] The study has been approved by our institute IRB.



(a) PIN number based authentication

(b) Multiple trails until success

(c) Gesture based authentication

Figure 14: Performance of verifying legitimate users when the testing region is below the vibration motor and receiver.

Credential-aware Attack. The attacker gets to know the legitimate user's PIN/lock pattern/gesture. But he has not observed how the legitimate user presses his/her PIN numbers or swipes his/her lock patterns and gestures on the authentication surface. The attacker performs the same PIN/lock pattern/gesture as the legitimate user did without knowing the legitimate user's detailed behavior. Each attacker inputs the PIN/lock pattern/gesture 10 times. In total, we collected 420 inputs.

Knowledgeable Observer Attack. The attacker not only knows the legitimate user's PIN/lock pattern/gesture but also observes how the legitimate user inputs them on the authentication surface. Each attacker practices 5 times and then inputs the PIN/lock pattern/gesture 10 times, trying to pass the authentication. Again, 420 inputs are collected.

Side-channel Attack. In addition, we perform the side-channel attack by placing additional vibration receivers on the authentication surface. In particular, two receivers are employed: one is placed adjacent to the original receiver, whereas the other is placed at the other side of the surface opposite to the original receiver.

8.3 Evaluation Metrics

Verification Accuracy/Attack Success Rate of PIN Number-based Authentication. The verification accuracy/attack success rate shows the percentage of correctly verified PIN numbers entered by the legitimate user or attacker respectively during the user authentication process. Specifically, it includes the complete PIN sequence verification accuracy and the PIN digit verification accuracy. The complete PIN sequence verification accuracy measures the rate of the user's input PINs being completely recognized (i.e., all numbers in the PIN sequence are correctly recognized), while the PIN digit identification accuracy shows the rate of successfully recognizing each single PIN digit.

Verification Accuracy/Attack Success Rate of Lock Patten-based Authentication. The verification accuracy/attack success rate shows the percentage of correctly verified lock patterns input by the legitimate user or attacker respectively during the user authentication phase. Similarly, it includes the complete lock pattern verification accuracy and lock pattern segment verification accuracy.

ROC Curve of Gesture-based Authentication. ROC curve is a plot of true positive rate (TPR) over false positive rate (FPR). The TPR denotes the rate of the legitimate users passing the authentication while FPR denotes the rate of the attackers successfully passing the system. Through varying the feature distance threshold in gesture-based authentication, we can achieve varied TPR and FPR and obtain ROC curves to evaluate the system performance.

8.4 System Performance of Verifying Legitimate Users

PIN Number-based Verification. Figure 14(a) shows the identification accuracy of each PIN digit and the complete PIN sequence of 15 legitimate users. Our PIN number based authentication can achieve a high verification accuracy. Specifically, the users can obtain over 95% verification accuracy of recognizing each PIN digit and the mean verification accuracy of the complete PIN sequence reaches 90%. Moreover, the verification accuracy of each PIN digit is higher than that of PIN sequence, since the complete PIN verification accuracy result requires that all the PIN numbers in the PIN sequence are correctly identified. The results demonstrate our system is effective in verifying all the legitimate users.

Lock Pattern-based Verification. Figure 14(b) shows the average authentication accuracy of the lock-pattern based verification with different number of trials. Specifically, the average verification accuracy of the complete lock pattern reaches 79% and 95% with a single trial or two trials respectively, which requires all the segments of the lock pattern to be correctly identified. In addition, the accuracy of the lock pattern identification is slightly lower than that of the PIN sequence based authentication, which indicates that swiping a finger continuously on the surface generates more errors than pressing the finger separately on each grid point. The above verification results show that our VibWrite can achieve a good performance to authenticate users by lock patterns.

Gesture-based Verification. Figure 14(c) illustrates the effectiveness of legitimate user verification in gesture-based authentication with ROC curves. 15 legitimate users perform their preferred simple gestures (i.e., one of our four predefined gestures as shown in Figure 10) ten times. With only one training instance (i.e., one time swiping) for each user, we observe that given a requirement of a 90% true positive rate, we can achieve as low as a 5% false positive rate on average, which indicates only around 5% of gesture trials have gained unauthorized access. We also observe that the using both DTW and EMD techniques can provide slightly better performance than that of only using EMD technique, since it considers the similarity in both time warped feature sequences and the features' distributions. The obtained high verification accuracy and the low-training efforts demonstrate that VibWrite is capable to distinguish different users even though they perform the same simple gesture due to their distinct behavioral biometrics (i.e., finger tip size and structures).

Multiple Authentication Trials and Fall-back Strategy. Figure 14(b) shows the average verification rate under different number of trials. We observe that our system can achieve over 99% verification rate with both of the PIN number and lock pattern inputs when users enter three trials. For the first-time user input, our system can achieve around 89% and 79% accuracies when users enter their PIN numbers or lock patterns, respectively. Additionally, our system can integrate with any fall-back strategy to let the legitimate user bypass the system, e.g., the legitimate user can always use a physical key to enter his vehicle/apartment.

8.5 Attacks on Legitimate User's Credentials

Under blind attacks, both our PIN number and lock pattern based authentications can achieve close to zero attack success rate. The results are intuitive because the attackers' random PIN guesses or lock pattern guesses are nearly impossible to pass the legitimate user's system within limited trials. Similarly, for gesture-based authentication, the TPR in the obtained ROC curve is close to 100% when the FPR is close to 0%, which shows that the attackers' random gestures cannot successfully access the system.

Under credential-aware attacks, our system also achieves high accuracy (i.e., close to 0% attack success rate) for all three types of authentications. Since the attackers do not possess the knowledge of the VibWrite setting details (e.g., grid size, gesture region and the authentication surface), the attackers' finger-inputs are hard to generate the similar impacts on the vibration propagation as the legitimate users do. Knowledgeable observer attack is the most extreme attack, where the attacker is capable of knowing the user's credentials and observing the legitimate user's finger inputs. Additionally, the attacker has the knowledge of the VibWrite setting details and can perform the finger inputs on the same authentication surface. Thus in the rest of this paper, we present the performance evaluation results of our system under this more challenging knowledgeable observer attack.

PIN Number-based Authentication. Figure 15(a) shows the performance of our VibWrite in PIN number based authentication under knowledgeable observer attack, where 1 of 15 users alternatively behaves as victim and other 14 users play as attackers. We find that the VibWrite system is very effective in defending against attackers even though they have the knowledge of the legitimate user's PIN and use the same VibWrite setting (e.g., grid size and authentication surface). In particular, the attackers can only break an average of around 7% single PIN digits. Furthermore, even if the attackers can successfully verify several PIN digits, it is even harder for them to break the complete PIN sequences of the legitimate user. In particular, the attackers can only achieve an average of 2% attack success rate in verifying complete PIN sequences.

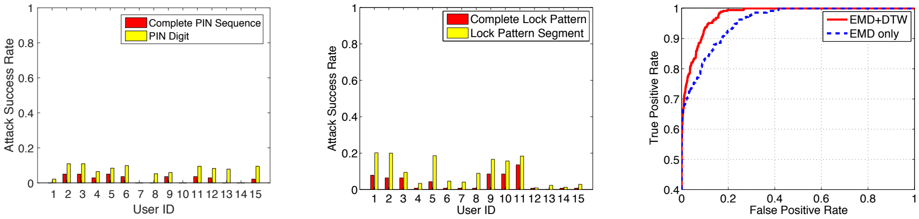
Lock Pattern-based Authentication. Similarly, we ask the 15 users to alternatively play one victim and fourteen attackers, who swipe 10 lock patterns after practice based on the knowledgeable observation. Figure 15(b) depicts the attack success rate of lock-pattern based authentication on each legitimate user under the knowledgeable observer attack. The results show that the attackers are hard to pass the system even though they imitate the legitimate user's behavior to swipe the same lock patterns on the same grid of the same authentication surface after practice. Specifically, for the user 4, 6-8 and 12-15, all the fourteen attackers can hardly pass the legitimate user's complete lock patterns in 10 trials though they can successfully swipe around 5% accurate segments of the lock patterns. The average attack success rates of the lock pattern segment and the complete lock pattern are around 5% and 11% respectively. Moreover, we find the performance of the lock pattern based authentication under knowledgeable observer attack is comparably good to that of the PIN number based authentication.

Gesture-based Authentication. We evaluate the performance of VibWrite in gesture-based authentication under knowledgeable observer attacks, where attackers try to mimic the legitimate user's swiping gestures. In order to test the worst case in VibWrite, we only rely on one single training data for the legitimate user. Figure 15(c) shows the ROC curve, where we can achieve as low as a 3% false positive on average given a requirement of a 80% true positive rate. Even for only using EMD technique, we can still achieve as low as a 8% false positive rate on average given a requirement of a 80% true positive rate. The results indicate that, even for the most challenging knowledgeable observer attack, VibWrite is still effective in defending against attackers and successfully authenticate legitimate users in the meanwhile.

8.6 Side-channel Attacks

Attacks via a Vibration Receiver. One may suspect that attackers can place hidden vibration receivers on the authentication surface to recover the vibration signals and obtain the unique features of the legitimate user. In reality, the hidden receiver cannot be placed at the exact same location as the VibWrite's receiver. Thus, our $Hidden1$ and $Hidden2$ are placed at two representative locations that an adversary may choose to launch a side-channel attack. Particularly, $Hidden1$ is placed adjacent to the original receiver, whereas $Hidden2$ is placed at the other side of the authentication surface (around $3cm$ thickness) opposite to the original receiver. Figure 16 shows the mean and standard deviation of the Pearson Correlation coefficients [8] between the signals received by the original receiver and two hidden receivers after the designed vibration chirps are generated 20 times. We observe that $Hidden1$ and $Hidden2$ can only achieve a very low correlation coefficient less than 0.2. This indicates that the vibration signals received by hidden receivers present very different patterns comparing to that received by the original receiver even when the hidden receivers are placed very close to the original receiver, making the attacks via a hidden vibration receiver ineffective.

Attacks via a Nearby Microphone. Furthermore, a nearby microphone can record the acoustic sounds emitted by the vibration motor, however, the additional transmission path (i.e., air between the vibration motor and microphone) can largely change the vibration patterns, making it also difficult to recover the similar vibration signals received by VibWrite's vibration receiver. Additionally, a few new studies demonstrate that physical vibrations can be recovered to a certain extent by using wireless signals [48]



(a) PIN number based authentication

(b) Lock pattern based authentication

(c) Gesture based authentication

Figure 15: Performance of user authentication under knowledgeable observer attacks when the testing region is below the vibration motor and receiver.

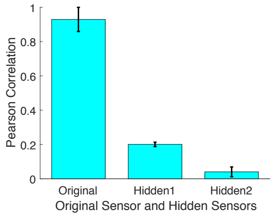


Figure 16: Similarity between the vibrations received by VibWrite's original receiver and hidden receivers.

and high-speed cameras [13]. However, these solutions can only recover relatively low-quality audio/vibrational signals due to the limits of the hardware sensing ability in both vibration amplitude and frequency. Thus, they are mainly used for eavesdropping human speech sounds whose frequency typically falls below $1KHz$.

8.7 Impact of Training Data Size

PIN Number/Lock Pattern based Verification. Our system can achieve around 90% accuracy in identifying each PIN digit/lock-pattern segment with the grid point training time over 0.4 seconds while the identification of complete PIN sequences or complete lock pattern achieve over 80% accuracy with the grid point training time over 0.6 seconds as shown in Figure 17. Moreover, the PIN sequence/lock pattern based authentication can achieve higher accuracy with longer training time and the accuracy reaches stable when the training size is over around 2 seconds.

Gesture-based Verification. From the results as shown in Figure 14(c) and Figure 15(c), we observe that our gesture-based verification can obtain very high authentication accuracy with the training profile only containing one single gesture training instance. The results also indicate that our gesture-based authentication system could work with a very small training data size.

8.8 Impact of Surface and Vibration Motor/Receiver Placement

We change the positions of the vibration motor and the piezoelectric sensor to the center of each side and evaluate the PIN sequence verification accuracy on the grid of the door panel surface. Ten users are first asked to construct their individual grid profile and then input their PIN sequences with this new experimental setup for verification. The results in Figure 18 show that our PIN number based authentication can achieve comparably high verification accuracy for this setup. In particular, the accuracies of verifying the complete PIN sequence and PIN digit are 88% and 94% respectively. The similar results can also be observed for lock pattern based and gesture based authentication. Thus our system is robust for different vibration generator/receiver placements.

9 DISCUSSION

Serving as a concrete starting point of vibration-based authentication system, VibWrite is a low-cost and easy-to-deploy solution that has a high potential to work at various places such as apartment buildings, hotel rooms, smart homes, etc. We admit that the current system is still not ready for the industrial deployment in terms of its authentication/false-accept rates, thus a large space is left for us to further improve the system. In this section, we introduce a few limitations of the current VibWrite system and the potential for future improvements.

Accuracy, and Further Improvement. The current system achieves around 89% and 79% authentication rates with a single trial when users enter their PIN numbers or lock patterns, respectively. The accuracy number is comparable to a few recent low-cost authentication/verification solutions (e.g., [23, 30, 47, 51]), which use either gait patterns captured by existing Wi-Fi/smartphone or passive sensing of embedded sensors on smartphones. Specifically, the gait pattern based solution could achieve around 80% detection rate of unauthorized users when leveraging accelerometers on smartphones [30] and 79% user recognition accuracy when using off-the-shelf Wi-Fi [47]. Multi-sensor (i.e., gyroscope, magnetometer and accelerometer) based smartphone authentication can achieve around 70% and 90% accuracy in the studies [51] and [23], respectively. However, the current VibWrite system is still far from practical deployment as a legitimate user may need to try a few times to pass the system. To improve the system performance, we target to explore the following aspects in our future work including deploying multiple sensor pairs, refining the hardware, and improving the authentication algorithms. Specifically, more than one pair of vibration transmitters and receivers can be employed to help increase the dimension of the surface sensing features, which can better represent each individual's behavioral and physiological characteristics. In addition, empirically we noticed that the uniqueness of the features is affected by the stableness of the hardware components as the weak analog signals extracted by the piezoelectric sensor can be easily distorted when passing through electronic components (e.g., amplifier and ADC). We thus could build a higher standard hardware signal processing component (e.g., ultra-low-noise signal amplifier) to enhance the system. Meanwhile, the improvement of the vibration motor in terms of its power level, stableness and frequency response could become another venue to explore.

Coping with Additional Physical Attacks. In addition to the side channel attacks via a hidden vibration receiver or a nearby microphone, other types of physical attacks might be launched when the system is deployed in practice. We discuss a couple of representative ones below and show how VibWrite could be extended in coping up with such attacks. Given that the proposed system is highly dependent on the attached surface, such surface dependency might be employed by an adversary to launch a denial-of-service (DoS) attack (e.g., adhering tiny objects or a hidden vibration motor to the surface) to prevent the legitimate user from passing the system. To combat the DoS attack, VibWrite can develop a simple mechanism to perform the surface sanity check periodically by comparing the received vibration signals with the *empty surface* training profile. If the surface dissimilarity is detected, the authentication surface will be examined. The most extreme case is when an adversary gets access to the cable connecting the vibration motor/sensor and cut it to make the system not function at all. On one hand, to deal with such a physical attack, the vibration motor and receiver could be placed at the opposite side of the authenticating surface hidden from the users and even placed inside some enclosed cases hard to access without authorization. On the other hand, the adversary does not gain much benefit in this attack as he still cannot pass the authentication system. We leave the detailed study of these adversarial cases as an avenue for our future work.

System Maintenance. As a starting point, our system is evaluated in a relatively stable indoor environment. However, in practical deployment, there are many environmental factors that need to be taken into consideration and may affect the system performance. For instance, if the surface (e.g., car door panel) is exposed to an outdoor environment, the surface's vibration response may be changed across different days affected by temperature, humidity, wind, wetness, dirt, etc. Additionally, the temporary presence of additional objects placed on the surface (e.g., a book placed on the desk) could alter the received vibrations slightly different from the trained one. The noticeable effect caused by these factors might be reduced through further filtering or directional sensing techniques. More robust machine learning methods grounded on deep learning [4] can also be built in our future work to deal with various environmental-related elements. In addition, future work should continue the evaluation with more/diverse population samples, longer time periods and more influential factors to improve the system robustness.

10 CONCLUSION

In this paper, we propose VibWrite, which implements the idea of low-cost low-power tangible user authentication beyond touch screens to any solid surface to support smart access applications (e.g., apartment entrances, vehicle doors, or smart appliances). Utilizing low-cost physical vibration, VibWrite performs ubiquitous user authentication via finger-input by integrating passcode, behavioral and physiological characteristics, and surface dependency together to provide enhanced security. VibWrite is built upon a vibration-based touch sensing technique that enables touching and writing on any solid surface through analyzing unique vibration signal features (e.g., frequency response and cepstral coefficient) in the frequency domain. It is easy to deploy and flexibly provides users with three independent forms of secrets (including PIN number, lock pattern, and simple gesture) to gain security access by developing new techniques of virtual grid point derivation, featured-based dynamic time warping (DTW) and distribution analysis based on earth mover's distance (EMD). We perform extensive experiments with participants input their passcodes by using three forms of secrets. We also study the robustness of Vibwrite under various attacks trying to impersonate the legitimate user or launching side-channel attacks to hack the VibWrite system directly. Our results indicate that VibWrite is resilient to side-channel attacks. And it can verify legitimate user with high accuracy under minimum training efforts while successfully deny the access requests from unauthorized users with a low false positive rate.

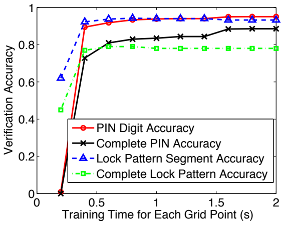


Figure 17: Performance of both PIN number based and lock pattern based authentications with different grid point training time periods.

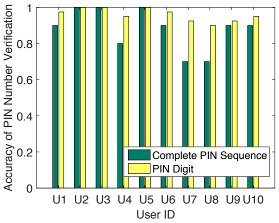


Figure 18: Performance of PIN number based authentication in verifying legitimate user when the testing region is on a door panel.

11 ACKNOWLEDGMENT

We thank our shepherd, Dr. Zhou Li, and the anonymous reviewers for their insightful feedbacks. This work was partially supported by the National Science Foundation Grants CNS-1514436, CNS-1716500 and CNS-1526524.

REFERENCES

[1] 2017. Access Control Market. http://www.marketsandmarkets.com/Market-Reports/access-control-market-164562182.html?gclid=CjwKEAjw9MrIBRCr2LPek5-h8U0SJAD3jfhtuAGYZKVZqHc8ZSphI146GhgExcOxXIts14fKyENFLRoCXG7w_wcB. (2017).
[2] 2017. Capacitive Sensor. http://www.sensorwiki.org/doku.php/sensors/capacitive. (2017).
[3] 2017. CR300. https://amgtime.com/hardware-biometric-fingerprint-reader-tm100. (2017).
[4] 2017. Deep Learning. http://www.ceva-dsp.com/app/deep-learning/. (2017).
[5] 2017. FacePass. https://crownsecurityproducts.com/facepass-facial-recognition-time-clock.html. (2017).
[6] 2017. Geovision. https://www.surveillance-video.com/security-84-fr20200-0010.html. (2017).
[7] 2017. How to implement fingerprint authentication in automobiles. http://www.electronicproducts.com/Sensors_and_Transducers/Sensors/How_to_implement_fingerprint_authentication_in_automobiles.aspx. (2017).
[8] 2017. Pearson Product moment Correlation Coefficient. http://en.wikipedia.org/wiki/Pearson-product-moment-correlation-coefficient. (2017).
[9] Arathi Arakala, Jason Jeffers, and Kathy J Horadam. 2007. Fuzzy extractors for minutiae-based fingerprint authentication. In *International Conference on Biometrics*. Springer, 760–769.
[10] Adam J Aviv, Katherine L Gibson, Evan Mossop, Matt Blaze, and Jonathan M Smith. 2010. Smudge Attacks on Smartphone Touch Screens. *Woot* 10 (2010), 1–7.
[11] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Issue 3. Software available at url-http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[12] 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.

[13] Abe Davis, Michael Rubinstein, Neal Wadhwa, Gautham J Mysore, Frédo Durand, and William T Freeman. 2014. The visual microphone: passive recovery of sound from video. *ACM Transactions on Graphics* (2014).

[14] Antonella De Angeli, Mike Coutts, Lynne Coventry, Graham I Johnson, David Cameron, and Martin H Fischer. 2002. VIP: a visual approach to user authentication. In *Proceedings of the working conference on advanced visual interfaces (ACM AVI)*. 316–323.

[15] Alexander De Luca, Alina Hang, Frederik Brudy, Christian Lindner, and Heinrich Hussmann. 2012. Touch me once and i know it's you!: implicit authentication based on touch screen patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 987–996.

[16] Rachna Dhamija and Adrian Perrig. 2000. Deja Vu-A User Study: Using Images for Authentication. In *USENIX Security Symposium*.

[17] RG Dong, AW Schopper, TW McDowell, DE Welcome, JZ Wu, WP Smutz, C Warren, and S Rakheja. 2004. Vibration energy absorption (VEA) in human fingers-hand-arm system. *Medical engineering & physics* 26, 6 (2004), 483–492.

[18] Benoit Duc, Stefan Fischer, and Josef Bigün. 1999. Face authentication with Gabor information on deformable graphs. *IEEE Transactions on Image Processing* 8, 4 (1999), 504–516.

[19] Richard O Duda, Peter E Hart, and David G Stork. 2012. *Pattern classification*. John Wiley & Sons.

[20] Alain Forget, Sonia Chiasson, and Robert Biddle. 2010. Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1107–1110.

[21] Pravein Govindan Kannan, Seshadri Padmanabha Venkatagiri, Mun Choon Chan, Akhihebbal L Ananda, and Li-Shiuan Peh. 2012. Low cost crowd counting using audio tones. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. 155–168.

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

[23] Wei-Han Lee and Ruby B Lee. 2015. Multi-sensor authentication to improve smartphone security. In *Information Systems Security and Privacy (ICISSP), 2015 International Conference on*. IEEE, 1–11.

[24] 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.

[25] Weizhi Meng, Wenjuan Li, Lijun Jiang, and Liying Meng. 2016. On Multiple Password Interference of Touch Screen Patterns and Text Passwords. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 4818–4822.

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

[27] K Sri Rama Murty and Bayya Yegnanarayana. 2006. Combining evidence from residual phase and MFCC features for speaker recognition. *IEEE Signal Processing Letters* 13, 1 (2006), 52–55.

[28] Phuc Nguyen, Ufuk Muncuk, Ashwin Ashok, Kaushik R Chowdhury, Marco Gruteser, and Tam Vu. 2016. Battery-Free Identification Token for Touch Sensing Devices. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. ACM, 109–122.

[29] Tomohiro Ohshima, Tomohiro Morita, Takeshi Tanaka, and Naotako Yamamoto. 2006. Indoor apparatus of intercom system and method for controlling indoor apparatus. (June 22 2006). US Patent App. 11/472,432.

[30] Yanzhi Ren, Yingying Chen, Mooi Choo Chuah, and Jie Yang. 2013. Smartphone based user verification leveraging gait recognition for mobile healthcare systems. In *Proceedings of IEEE SECON*.

[31] Yanzhi Ren, Yingying Chen, Mooi Choo Chuah, and Jie Yang. 2015. User Verification Leveraging Gait Recognition for Smartphone Enabled Mobile Healthcare Systems. *IEEE Transactions on Mobile Computing* 14, 9 (2015), 1961–1974.

[32] Yanzhi Ren, Chen Wang, Yingying Chen, Mooi Choo Chuah, and Jie Yang. 2015. Critical segment based real-time E-signature for securing mobile transactions. In *Proceedings of IEEE Conference on Communications and Network Security (CNS)*. 7–15.

[33] Kenneth Revett. 2009. A bioinformatics based approach to user authentication via keystroke dynamics. *International Journal of Control, Automation and Systems* 7, 1 (2009), 7–15.

[34] Manning I Rose and Lee W Hoevel. 1998. Access card for multiple accounts. (June 23 1998). US Patent 5,770,843.

[35] Y. Rubner and Stanford University. Computer Science Dept. 1999. *Perceptual metrics for image database navigation*. Number 1621 in Report STAN-CS-TR. Stanford University. http://books.google.com/books?id=5b1EAQAAIAAJ

[36] Michael Sherman, Gradeigh Clark, Yulong Yang, Shridatt Sugrim, Arttu Modig, Janne Lindqvist, Antti Oulasvirta, and Teemu Roos. 2014. User-generated free-form gestures for authentication: Security and memorability. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. ACM, 176–189.

[37] Jaspreet Singh, Upamanyu Madhow, Rajesh Kumar, Subhash Suri, and Richard Cagley. 2007. Tracking multiple targets using binary proximity sensors. In *Proceedings of the 6th international conference on Information processing in sensor networks*. ACM, 529–538.

[38] Bernard Sklar. 2001. *Digital communications*. Vol. 2. Prentice Hall NJ.

[39] Joshua R Smith, Kenneth P Fishkin, Bing Jiang, Alexander Mamishev, Matthai Philipose, Adam D Rea, Sumit Roy, and Kishore Sundara-Rajan. 2005. RFID-based techniques for human-activity detection. *Commun. ACM* 48, 9 (2005), 39–44.

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

[41] Stephen P. Tarzia, Peter A. Dinda, Robert P. Dick, and Gokhan Memik. 2011. Indoor localization without infrastructure using the acoustic background spectrum. In *Proceedings of the 9th international conference on Mobile systems, applications, and services (ACM MobiSys)*.

[42] Gineke A ten Holt, Marcel JT Reinders, and EA Hendriks. 2007. Multi-dimensional dynamic time warping for gesture recognition. In *Thirteenth annual conference of the Advanced School for Computing and Imaging*, Vol. 300.

[43] Jing Tian, Chengzhang Qu, Wenyuan Xu, and Song Wang. 2013. KinWrite: Handwriting-Based Authentication Using Kinect.. In *NDSS*.

[44] Adam T Timmons and Osman D Altan. 2010. Grid unlock. (Feb. 2 2010). US Patent App. 12/698,321.

[45] Yu-Chih Tung and Kang G Shin. 2016. Expansion of human-phone interface by sensing structure-borne sound propagation. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 277–289.

[46] Tam Vu, Akash Baid, Simon Gao, Marco Gruteser, Richard Howard, Janne Lindqvist, Predrag Spasojevic, and Jeffrey Walling. 2012. Distinguishing users with capacitive touch communication. In *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 197–208.

[47] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 363–373.

[48] Teng Wei, Shu Wang, Anfu Zhou, and Xinyu Zhang. 2015. Acoustic eavesdropping through wireless vibrometry. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 130–141.

[49] Susan Wiedenbeck, Jim Waters, Leonardo Sobrado, and Jean-Camille Birget. 2006. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *Proceedings of the working conference on Advanced visual interfaces*. ACM, 177–184.

[50] Nan Zheng, Aaron Paloski, and Haining Wang. 2011. An efficient user verification system via mouse movements. In *Proceedings of the 18th ACM conference on Computer and communications security*. ACM, 139–150.

[51] Jiang Zhu, Pang Wu, Xiao Wang, and Joy Zhang. 2013. Sensec: Mobile security through passive sensing. In *Computing, Networking and Communications (ICNC), 2013 International Conference on*. IEEE, 1128–1133.