9-2018

BiLock: User authentication via dental occlusion biometrics

Yongpan ZOU

Meng ZHAO

Zimu ZHOU
Singapore Management University, zimuzhou@smu.edu.sg

Jiawei LIN

Mo LI

See next page for additional authors

Citation

ZOU, Yongpan; ZHAO, Meng; ZHOU, Zimu; LIN, Jiawei; LI, Mo; and WU, Kaishun. BiLock: User authentication via dental occlusion biometrics. (2018). *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2, (3), 152:1-20. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/4692

BiLock: User Authentication via Dental Occlusion Biometrics

YONGPAN ZOU, College of Computer Science and Software Engineering, Shenzhen University
MENG ZHAO, College of Computer Science and Software Engineering, Shenzhen University
ZIMU ZHOU, Computer Engineering and Networks Laboratory, ETH Zurich
JIAWEI LIN, College of Computer Science and Software Engineering, Shenzhen University
MO LI, School of Computer Science and Engineering, Nanyang Technological University
KAISHUN WU, College of Computer Science and Software Engineering, Shenzhen University

User authentication on smart devices is indispensable to keep data privacy and security. It is especially significant for emerging wearable devices such as smartwatches considering data sensitivity in them. However, conventional authentication methods are not applicable for wearables due to constraints of size and hardware, which makes present wearable devices lack convenient, secure and low-cost authentication schemes. To tackle this problem, we reveal a novel biometric authentication mechanism which makes use of sounds of human dental occlusion (*i.e.*, tooth click). We demonstrate its feasibility by comprehensive measurement study, and design a prototype-BiLock with two Android platforms. Extensive real-world experiments have been conducted to evaluate the accuracy, robustness and security of BiLock in different environments. The results show that BiLock can achieve less than 5% average false reject rate and 0.95% average false accept rate even in a noisy environment. Comparative experiments also demonstrate that BiLock possesses advantages in robustness to noise and security against *replay* and *observation attacks* over existing voiceprinting schemes.

CCS Concepts: • **Human-centered computing** → **Ubiquitous computing**; *Mobile computing*; *Mobile devices*; • **Security and privacy** → Privacy protections;

Additional Key Words and Phrases: Dental occlusion; Biometric authentication; Mobile devices

ACM Reference Format:
Yongpan Zou, Meng Zhao, Zimu Zhou, Jiawei Lin, Mo Li, and Kaishun Wu. 2018. BiLock: User Authentication via Dental Occlusion Biometrics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 152 (September 2018), 20 pages.
https://doi.org/10.1145/3264962

1 INTRODUCTION

Biometric authentication on personal devices such as smartphones and smartwatches is becoming increasingly popular. These small form-factor devices contain enormous amounts of sensitive data *e.g.* contacts, messages and health records stored, and frequently used in public places. Conventional authentication mechanisms *e.g.* passwords may involve cumbersome input efforts on small devices and are more vulnerable to shoulder surfing and smudge attacks [33]. Alternatively, biometric authentication such as fingerprints [22, 23], face/iris

Authors' addresses: Yongpan Zou, College of Computer Science and Software Engineering, Shenzhen University, yongpan@szu.edu.cn; Meng Zhao, College of Computer Science and Software Engineering, Shenzhen University, zhaomeng.szu.edu@gmail.com; Zimu Zhou, Computer Engineering and Networks Laboratory, ETH Zurich, zzhou@tik.ee.ethz.ch; Jiawei Lin, College of Computer Science and Software Engineering, Shenzhen University, lin364884292@foxmail.com; Mo Li, School of Computer Science and Engineering, Nanyang Technological University, limo@ntu.edu.sg; Kaishun Wu, College of Computer Science and Software Engineering, Shenzhen University, wu@szu.edu.cn.

2474-9567/2018/9-ART152 $15.00
https://doi.org/10.1145/3264962

Fig. 1. Application scenarios of BiLock. A user places his/her smartphone or smartwatch 5 cm to 15 cm away to his/her lips, and performs an occlusion gesture. The microphone embedded in the smartphone or smartwatch records the sounds generated by the occlusion gesture and extracts features, which are fed into a classifier for user authentication. BiLock is able to achieve average FRR and FAR of 4.33% and 0.82% in presence of noise up to 50 $\sim$ 60 dB (measured next to the device) in indoor environments.

recognition [11, 46], and voice-prints [8, 17], are attracting extensive research interests and resulting in commercial successes, for their more user-friendly and secure usage [4].

Designing a ubiquitous, robust and socially acceptable biometric authentication mechanism for small form-factor devices *e.g.* smartwatches is non-trivial. For example, fingerprints require a capacitive or ultrasonic sensor to function [29], whose size is prohibitive to embed into wearable devices such as smartwatches. Face/iris recognition involves additional hardware and intensive computation not always affordable on resource-constrained devices [13]. Due to the ubiquity of microphones and the lower energy footprints (compared with face recognition), voice-based schemes have been increasingly applied in smartphone and smartwatch authentication, such as Google Pixel and Samsung Gear Live smartwatch [31]. Yet they are still considered annoying in libraries and offices and susceptible to changes (*e.g.* when the user caught a cold [30]).

In this paper, we propose BiLock, a new biometric authentication scheme for small personal devices based on *dental occlusion*. In principle, BiLock extracts unique signatures from the sounds generated by a user's occlusion activities, which are recorded by the built-in microphone of a smartphone or a smartwatch placed close to the user's lips (see Fig. 1). BiLock brings about the following advantages. *(i) Ubiquitous.* Microphones are available in a wide range of mobile, wearable and IoT devices. *(ii) Socially acceptable.* Dental occlusion acoustics is more socially acceptable than voice-based schemes in public places such as libraries and offices, which is because sounds of tooth clicks are more imperceptible and unobtrusive to others. It has also be exploited as a convenient hand-free interface in the HCI community [3][37]. *(iii) Robust.* Sounds generated by dental occlusion are resilient to human speeches and other interference due to their differences in frequencies. *(iv) Reliable.* Using dental occlusion acoustics for authentication is more reliable than touch-based methods such as PIN and lock patterns in mobile scenarios where they are likely to produce input errors.

Contributions and Roadmap. To enable a working system, we need to answer the following problems. *(i) Are the sounds generated by dental occlusion distinctive for different users while consistent for the same user over time? (ii) How to extract signatures from dental occlusion sounds for accurate and robust authentication?* The design of BiLock addresses the above challenges with the following contributions.

- Using measurements of 50 participants collected over 6 months, we demonstrate the diversity of dental occlusion sounds captured by commodity embedded microphones across users, and their consistency for the same user. The measurements serve as an empirical feasibility study of dental occlusion acoustics as a new biometric modality for authentication or access control for tens of users.

- We design a pipeline that segments raw dental occlusion sounds, extracts effective features, and accurately classifies the generated signatures for authentication.
- We prototype BiLock with a smartphone and a smartwatch, and comprehensively evaluate its performance in various contexts. Results show that BiLock achieves a false reject rate (FRR) of 4.33% and a false accept rate (FAR) of 0.82% even in presence noise of $50 \sim 60\,dB$ in indoor environments. We also demonstrate it is more robust and secure than two commercial voice-based authentication schemes.

In the rest of this paper, we review related work in Sec. 2, present the measurement study in Sec. 3 and elaborate on the design of BiLock in Sec. 4. We show the evaluations in Sec. 5, discuss the future work inSec. 6 and conclude this paper in Sec. 7.

2 RELATED WORK

BiLock is related to the following categories of research.

2.1 Biometric Authentication on Mobile and Wearable Devices

Examples of physiological and behavioral characteristics for authentication include fingerprints, voice, iris, retina, face, keystrokes *etc.* and new measures are under development all the time [43]. To perform authentication on mobile and wearable devices, various biometrics measurable by different sensors on mobile and wearable devices have been explored. Fingerprints are widely adopted in current mobile devices, which are recognized by capacitive sensors. However, fingerprints may fail in case of water, cuts or bruises on fingers [20]. They also require a sensor sized of a fingertip, which may not always be available on wearable devices such as smartwatches. The development of computer vision has enabled face recognition such as Apple's FaceID [2] as an attractive authentication mechanism. Yet the computation overhead still prohibits its usage on resource-constrained wearable devices. Keystrokes and finger gestures are prevailing for user authentication on devices with a touch-screen [35][12][27][28]. These efforts exploit motion sensors *e.g.* accelerometers to extract the unique patterns in the rhythm, strength and other attributes of writing/tapping behaviours. However, the performance of these systems are vulnerable to user movements such as walking. In contrast, gait-based authentication is performed during walking using inertial sensors in smartphones [19]. Other biometrics such as heart rate [45] and brain waves [9] have also been proposed due to the popularity of medical wearable devices. BreathPrint [7] is a recent work that utilizes the sounds of breath detected by phone microphones to verify legitimate users.

As concluded in [43], there is no single best biometric feature for authentication in terms of robustness, distinctiveness, availability, accessibility and acceptability. Our work aims to explore dental occlusion as a new biometric authentication mechanism for mobile and wearable devices equipped with commodity microphones. Compared with other features detectable by microphones, such as voice and breath, sound of dental occlusion enjoy the advantages of being more secure to replay attack, more robust to external interference, and more stable to users' physiological status.

2.2 Dentistry and Tooth Click Interfaces

Sounds of dental occlusion have long be utilized in dentistry. Stewart *et al.* [39] first proposed to diagnose occlusion problems using the sounds of tooth clicks. Such diagnosis was conducted by professional dentists in the early days and has been digitalized with the development of computer technologies [36][32]. In addition to gnathosonic research, tooth clicks have also been exploited in assistive HCI interface designs for the disabled. Researchers have proposed tooth click based input or control interfaces by recording the sounds of tooth clicks with bone-conduction microphones [24][25] or by capturing tooth click induced vibrations with accelerometers [37][38]. Furthermore, some recent work [3] tries to localize the pair of teeth clicks where the sounds of tooth clicks come for hands-free interaction. BiLock shares the similar principle of gnathosonic research. However, BiLock

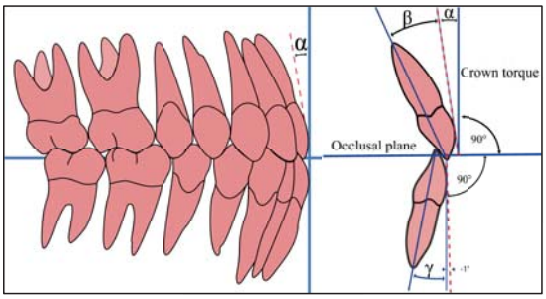


Fig. 2. An illustration of dental occlusion. Diversity in dental occlusion among individuals results in different sounds of dental occlusion [21].

harnesses the sound of tooth clicks as a biometric feature for user authentication, rather than an indicator of normal or abnormal occlusion. BiLock can be combined with previous research on tooth click interfaces once the user is authenticated.

2.3 Sensing Non-speech Body Sounds

Many research efforts have employed microphones in mobile and wearable devices to perceive non-speech body sounds for activity recognition. For instance, BodyScope [44] is a wearable acoustic sensing system attached to a user's neck to recognize activities such as eating, drinking, and coughing. SymDetector [40] is a smartphone-based application to detect sound-related respiratory symptoms such as sneeze, cough and sniffle. The microphone of smartphones have also been used to detect body sounds during sleep to assess sleep quality [15] or infer sleep stage [14]. All these systems can unobtrusively recognize activities that produce non-speech body sounds with a high accuracy. BiLock also senses non-speech sounds (from tooth clicks in our case). But instead of exploiting these sounds for activity recognition, BiLock takes these sounds as unique signatures for user authentication.

3 MEASUREMENTS AND FEASIBILITY ANALYSIS

This section presents a measurement study to show that the dental occlusion characteristics are different across users while consistent for the same user.

3.1 Physiological Mechanism of Dental Occlusion

Dental occlusion refers to the process of contacting between the maxillary (upper) and mandibular (lower) teeth. In this work, we make use of the *sounds* produced during dental occlusion for authentication. Such sounds are detectable by commodity microphones and differ for individuals who have different patterns of dental occlusion, which is the rationale of our paper. The diversity of dental occlusion comes from the differences in the shapes, sizes and the positions of teeth, which has been shown in medical studies such as [21, 34, 42]. Fig. 2 illustrates the parameters that depict the differences in dental occlusion.

3.2 Measurements

We collected the sounds of dental occlusion of 50 volunteers (labeled as $V_1 \sim V_{50}$, 30 males and 20 females, aged 20 to 43). The sounds were measured by the microphone embedded in a Samsung Galaxy Tab S2 tablet and a Huawei Watch 2. The smartwatch contains a Quad-core 1.1 GHz CPU and 768 M RAM, and runs Android Wear OS 2.0 operating system. The Samsung tablet consists of a Qualcomm Snapdragon 652 CPU and 3 GB RAM, and runs an Android 6.0 operating system. Both platforms have a single microphone sensor without noise reduction.

Fig. 3. Setup of measurement study

We developed an Android application to record the sounds and converted raw audio signals into 16 bit linear *.wav* files. We set a sampling rate of 44.1 kHz on the two devices.

The measurements were collected in two indoor environments, a meeting room and a lab room. In the meeting room, we artificially generated noises of 4 average levels by controlling the volume of a TV, *i.e.* 30 $\sim$ 40 dB (N_1), 40 $\sim$ 50 dB (N_2), 50 $\sim$ 60 dB (N_3) and 60 $\sim$ 70 dB (N_4), respectively. In the lab room, the average noise level was 50 $\sim$ 60 dB. In the meeting room, the participants remained seated (see the middle sub-figure in Fig. 3), while in the lab room, the participants walked in straight lines while performing the tooth click gestures (see the right sub-figure in Fig. 3). Thus there were a total number of 5 settings for the measurement study.

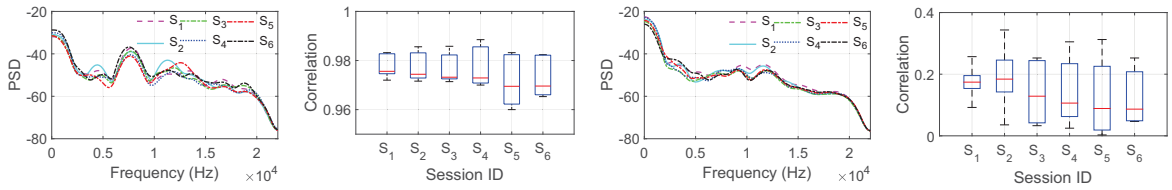
Before data collection, we explained to each participant the purpose of the project, the principle of BiLock, and its usage. We then asked each participant to experiment with tooth clicks (on both sides) until he/she found the way he/she could comfortably perform the gesture repeatedly. Most participants were able to figure out a comfortable way within 3 $\sim$ 5 clicks. Then in each setting, each participant put the tablet or smartwatch about 10 cm away from his/her lip, and performed tooth clicks for 100 times during 6 sessions. To make experiments closer to practical scenarios, we asked the participants to perform tooth clicks naturally as they do in daily life, and have a short rest every 5 continuous clicks. The sounds of every 5 tooth clicks are recorded at a time and stored in a single *.wav* file. We collected data from different sessions to investigate the variations in the tooth click gestures of the same user over time. Specifically, we regarded the first day of the measurement as Session 1 (S_1). The subsequent sessions were 3 $\sim$ 4 days (Session 2, S_2), 2 $\sim$ 3 weeks (Session 3, S_3), 1 $\sim$ 2 months (Session 4, S_4), 3 $\sim$ 4 months (Session 5, S_5), and 5 $\sim$ 6 months (Session 6, S_6) after the first day, respectively. The number of repeated instances was 20 for each during $S_1 \sim S_4$ and 10 for each during $S_5 \sim S_6$. We restricted the number of repetitions in each session to avoid fatigue. In summary, we collected a total number of 100 (number of instances) $\times$ 5 (number of settings) $\times$ 50 (number of participants) = 25,000 instances in the measurement study.

3.3 Data Analysis

In this section, we shall conduct an analysis with the dataset collected by the above experiments, in order to demonstrate the following observations that we make.

- *The same person show consistent signal patterns of tooth click over time.*
- *Different people show different signal patterns when they perform tooth click.*

3.3.1 Intra-User Analysis. We use samples of a participant X collected in each of the six sessions ($S_1 \sim S_6$) under 30 $\sim$ 40 db noise in the meeting room, and compute their average power spectrum density (PSD) in each session as shown in Fig. 4(a). Intuitively, we can see that the average PSD of collected samples remain highly consistent over different sessions. To quantify this, we calculate the correlation coefficients of PSD curves



(a) The PSD curves of participant X in different sessions (b) The PSD correlation coefficients of user X (c) The PSD curves of participant Y in different sessions (d) The PSD correlation coefficients between user X and user Y

Fig. 4. The PSD curves and correlation coefficients of samples collected in different sessions from different participants

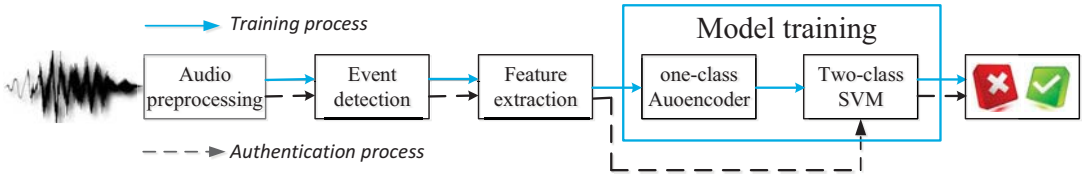


Fig. 5. The overall data processing flowchart of BiLock

between the first session (*i.e.*, S_1) and other sessions (*i.e.*, $S_1 \sim S_6$), and obtain the corresponding results as shown in Fig. 4(b). As we can see, the average correlation coefficients between different sessions are all above 0.96 even after a half year, which validates that sound of dental occlusion is a favorable biometric for user authentication. However, we can also notice that although the correlation keeps high, it decreases slightly over time as whole. The underlying reason is that people may slightly change the way of clicking tooth over a relatively long time.

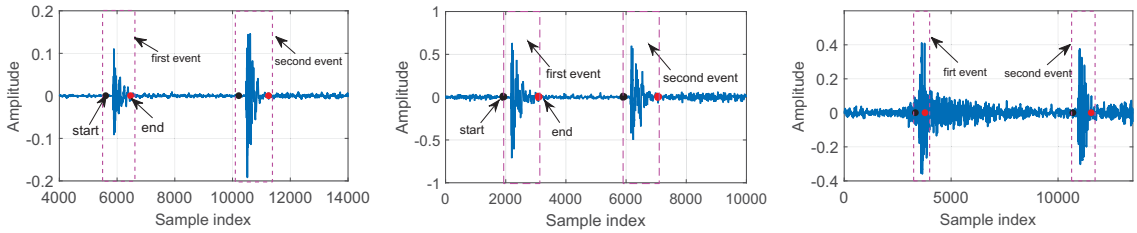
3.3.2 Inter-User Analysis. On the other hand, we also demonstrate the uniqueness of this novel biometric by plotting the PSD curves of audio samples collected from another participant Y as shown in Fig. 4(c). Comparing with Fig. 4(a), the overall trend of PSD curves computed for participant Y exhibits noticeable difference from those of user X. Correspondingly, we also compute the correlation coefficients of PSD curves calculated for user X and Y in different sessions as shown in Fig. 4(d). We can see that the overall correlation coefficients between different users are below 0.5, which is obviously lower than those of the same person in Fig. 4(b). The results indicate that sounds of dental occlusion show noticeable difference among different people.

4 BILOCK DESIGN

This section presents the design of BiLock. At a high level, BiLock processes raw audio signals, detects occlusion events, extracts features from segments containing the occlusion events, and feeds them into a two-class classifier to determine whether the signal segments come from a legitimate user or not (Fig. 5). As next, we elaborate on how to properly adopt acoustic signal processing and machine learning techniques to enable a functional occlusion based user authentication pipeline.

4.1 Audio Preprocessing

As with other non-speech body sounds, the sounds of dental occlusion are subtle and vulnerable to other interference such as human speech. Therefore the raw audio signals need to be filtered to enhance the signal to noise radio (SNR). Fig. 4 shows the average power spectral density of 100 instances of dental occlusion gestures collected in the measurement (see Sec. 3.2). As is shown, over 90% of the power spectrum distributes within



(a) Events detection under 40~50 db noise (b) Events detection under 50~60 db noise (c) Events detection under 60~70 db noise

Fig. 6. Examples of occlusion event detection under noise levels of (a) 40 ~ 50 dB, (b) 50 ~ 60 db (c) 60 ~ 70 dB

[10, 15000] Hz. Hence we apply a 6-order Butterworth filter with a passband of [10, 15000] Hz to filter out-of-band interference. Since the frequency band of dental occlusion overlaps with that of human speech, we further utilize wavelet de-noising to improve SNR in the passband. Wavelet de-noising is suited in analyzing short-duration, transient and abrupt signals and has been adopted in other non-speech body sound sensing schemes such as [1, 41]. In BiLock, we use a 3-level discrete wavelet decomposition with Daubechies 3 (db3) wavelet as mother wavelet, followed by soft coefficient thresholding. We refer readers to [6, 10] for details on wavelet de-noising.

4.2 Occlusion Event Detection

After de-noising, the next step is to detect dental occlusion events and extract the corresponding signal segments. BiLock adopts an adaptive energy-based event detection scheme to deal with a wide range of noise levels. Specifically, we apply a sliding window of width W on the signal $S(i)$ (i is the sample index). Assume the power of remaining noise follows a Gaussian distribution and denote the mean and the standard deviation of signal power at index i are $\mu(i)$ and $\sigma(i)$, respectively. Then the average signal power within a sliding window is calculated by:

$$\begin{aligned} \mu(i) &= \frac{1}{W}A(i) + (1 - \frac{1}{W})\mu(i-1) \\ \sigma(i) &= \frac{1}{W}B(i) + (1 - \frac{1}{W})\sigma(i-1) \end{aligned} \tag{1}$$

where $\mu(0) = 0$ and $\sigma(0) = 0$. $A(i)$ and $B(i)$ represent the cumulated power and the overall standard deviation of signals within a sliding window, respectively, where

$$\begin{aligned} A(i) &= \frac{1}{W}\sum_{k=i}^{W+i}|S(k)|^2 \\ B(i) &= \sqrt{\frac{1}{W}\sum_{k=i}^{W+i}(|S(k)|^2 - A(k))^2} \end{aligned} \tag{2}$$

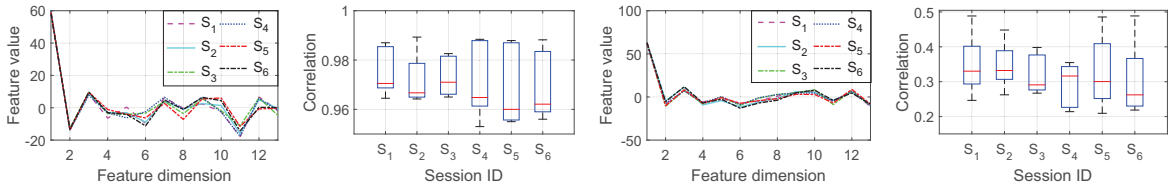
Finally a potential start point of $S(i)$ can be determined if

$$|S(i)|^2 > \mu(i) + \gamma_1\sigma(i) \tag{3}$$

where γ_1 is a constant independent of the noise level. Similarly, an end point of $S(i)$ is detected if

$$|S(i)|^2 < \gamma_2\overline{\mu} \tag{4}$$

where γ_2 is also a constant independent of the noise level. $\overline{\mu}$ is the average noise power when there is no dental occlusion event. In BiLock, we empirically set W, γ_1 and γ_2 as 6 ms, 2 and 5, respectively. We further examine the



(a) The average feature vector of user X in different sessions (b) The feature vector correlation coefficients of user X (c) The average feature vector of user Y in different sessions (d) The feature correlation coefficients of user X and user Y

Fig. 7. The feature vectors and correlation coefficients of samples collected in different sessions from different users

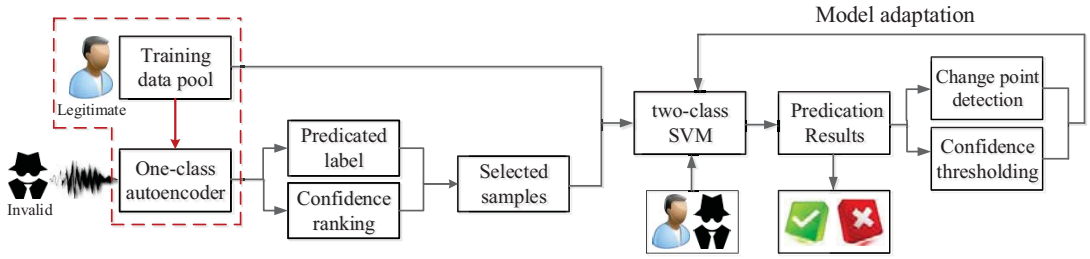


Fig. 8. The specific illustration of the model training and adaptation process

amplitudes of the detected peaks and remove those peaks with amplitude below 0.05 to avoid fake peaks. Fig. 6 shows two examples of event detection and segmentation with noise levels of 40 ∼ 50 db, 50 ∼ 60 db and 60 ∼ 70 dB, respectively.

4.3 Feature Extraction

Distinctive and robust features are crucial for an authentication system. In BiLock, we extract 13-order Mel-frequency Cepstral coefficients (MFCCs) from each signal segment with a Hamming window. We do not perform framing here because the duration of each event is relatively short (about 30 ms). Fig. 7 plots the correlation of occlusion sounds represented by the 13-order MFCCs of the same person and across persons, respectively. As is shown, the occlusion sounds represented by MFCCs still correlate for the same person, and are distinctive for different persons. The correlations are similar to those calculated from the raw PSDs.

4.4 Model for Authentication

As with previous user authentication schemes [5][35], we adopt a support vector machine (SVM) as the classifier for user authentication. Specifically, we use a two-class SVM classifier with radial basis function (RBF) kernel in BiLock. In its practical usage, a user authentication system may collect samples from both legitimate users and impostors. As shown in Fig. 8, we design techniques to harness these samples in model training (Sec. 4.4.1) and model adaptation (Sec. 4.4.2), as described below.

4.4.1 Model Training. Unlike previous studies on acoustic based activity recognition [14][15][40][44], a user authentication system may only have a limited number of positively (*i.e.* legitimate) labelled samples and some unlabelled samples (can be both positive and negative) for training. For instance, a voice-print authentication application may collect a few samples from the legitimate user to initialize, and some unlabelled samples to further train its classifier. Therefore, a challenge for BiLock is how to train a *two-class* classifier with limited

amounts of positively labelled training samples. We bootstrap the training of the two-class SVM classifier by *(i)* constructing a one-class auto-encoder to assign pseudo-labels to the unlabelled training samples and *(ii)* train the SVM using pseudo-labelled samples with high confidence.

We first construct a one-class auto-encoder using the labelled samples from the legitimate user. We use an auto-encoder because it often outperforms simpler classifiers such as SVMs. An auto-encoder is a special neural network with an input layer, multiple hidden layers and an output layer. It learns a latent and compressed representation of input data and then reconstructs output from this presentation, aiming to minimize the difference between the input and the output. The learned latent representation is called a *coder*. The process of obtaining the coder from the input is called *encoding* and reconstructing output from the coder is called *decoding*. Denote $\mathcal{I}$, $\mathcal{F}$ and $\hat{\mathcal{I}}$ as the input, representation and output domains, respectively. Further use ϕ and ψ to represent the encoding ant the decoding process. Then an auto-encoder $\mathbb{A} = (\mathcal{I}, \phi, \psi)$ works as follows.

$$\phi : \mathcal{I} \rightarrow \mathcal{F} \tag{5}$$

$$\psi : \mathcal{F} \rightarrow \hat{\mathcal{I}} \tag{6}$$

$$\phi, \psi = \arg\min_{\phi,\psi} ||\boldsymbol{I} - \hat{\boldsymbol{I}}||^2 = \arg\min_{\phi,\psi} ||\boldsymbol{I} - (\phi \circ \psi)\boldsymbol{I}||^2 \tag{7}$$

where $\boldsymbol{I}$ and $\hat{\boldsymbol{I}}$ are the input and output data sets, respectively. Since the number of nodes in the hidden layers are usually smaller than that of the input layer, the obtained representation can be viewed as a compressed version of the input data. In BiLock, we use an auto-encoder with one input layer with 13 nodes, one hidden layer with 8 nodes, and one output layer with 13 nodes. The transfer functions of encoder and decoder are *saturating linear function* and *pure linear function*, respectively. We set a threshold τ as the maximum reconstruction error of the positively labelled training samples.

Then for each unlabelled sample $\boldsymbol{x}$ in the training data, we feed it into the auto-encoder to calculate its reconstruction error Δ. If Δ is smaller than the threshold τ, then the sample is pseudo-labelled as positive (legitimate). We then sort the samples based on their reconstruction error, and select κ positive/negative samples with the smallest/largest reconstruction errors to be fed into the two-class SVM for training. We empirically set κ as 60% in our system implementation according to evaluation results in Sec. 5.2.

4.4.2 Model Adaptation. As shown in Sec. 3.3, the correlation of dental occlusion sounds of the same person still decreases gracefully over time. Therefore the model for authentication also needs to be updated after a period of time. To make the model up-to-date, we propose a model adaptation scheme as follows.

A sample to be tested is first fed into the two-class SVM classifier, which is labeled as either positive or negative. The sample is also associated with a confidence given by the SVM model. If the confidence exceeds a threshold $\beta = 0.5$ and the sample substantially differs those in the training set, the sample will be added into the training set to retrain the model at a later time. We utilize Kullback-Leibler (KL) divergence as a metric to quantify the difference between two feature vectors $\overline{MFCC_i}^t$ $(i = 1, 2, .., M)$ and $MFCC_i^{t+\Delta t}$ $(i = 1, 2, .., M)$ as follows:

$$KL(\Delta t) = \sum_{i=1}^{M} \overline{MFCC_i}^t \, log \frac{\overline{MFCC_i}^t}{MFCC_i^{t+\Delta t}} \tag{8}$$

where $M = 13$ in our case, $\overline{MFCC_i}^t$ $(i = 1, 2, .., M)$ is the mean feature vector of the samples in the training set collected up to time t, $MFCC_i^{t+\Delta t}$ $(i = 1, 2, .., M)$ is the feature vector of a newly sample tested at time $t + \Delta t$. By empirically setting a threshold η for the KL divergence, BiLock detects significantly different samples and uses them to update the SVM model.

5 EVALUATION

The section presents the evaluations of BiLock. We first introduce the evaluation setups in Sec. 5.1, then show the overall performance of BiLock in Sec. 5.2, compare the security of BiLock with two commercial voice-printing schemes in different contexts in Sec. 5.4, and finally investigate user experience of BiLock in Sec. 5.5.

5.1 Experiment Setup

As aforementioned, we implement BiLock as an Android App with Java on two platforms. The training of classifiers are conducted on a desktop with four-core Intel (R) Xeon (R) E3-1231 CPU and 16 G RAM running Windows 8 with MatLab R2015b software. The trained classifiers, as well as audio processing and event detection algorithms are fed to a Samsung Galaxy Tab S2 and a Huawei Watch 2, which are used for the measurement study in Sec. 3.

All the evaluations in Sec. 5.2 and Sec. 5.3 are performed on datasets collected in the measurement study (see Sec. 3.2) unless specified otherwise, with the following metrics.

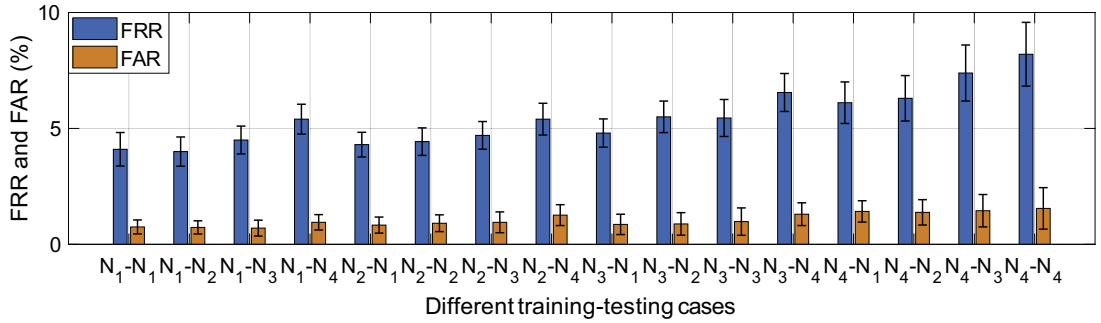
- False Accept Rate (FAR): FAR is a measure of likelihood that an authentication system incorrectly accepts an access trial of an unauthorized user.
- False Reject Rate (FRR): FRR represents the likelihood that an authentication system incorrectly rejects the access attempt of an legitimate user.

In the evaluation, we treat each volunteer (*e.g.*, V_1) as a valid user and other volunteers (*e.g.*, $V_2 \sim V_{50}$) as impostors. With a small number of samples (*i.e.*, 20 for the valid user without specification) from the valid user and impostors, we train an authentication model following the routine as described in Sec. 4. After that, we can obtain a model that can accept V_1 as a valid user and reject $V_2 \sim V_{50}$. We then test the model by feeding all remaining samples of $V_1 \sim V_{50}$. When a testing sample of the impostors (*i.e.*, $V_2 \sim V_{50}$) is wrongly accepted, a false accept instance occurs; when a testing sample of the valid user (*i.e.*, V_1) is wrongly rejected, a false reject instance occurs. By calculating the ratios false accept/reject instances, we can obtain the FAR/FRR of BiLock. For each valid user case, we repeat the above process for 10 times by randomly selecting the training samples and compute the average FAR and FRR.

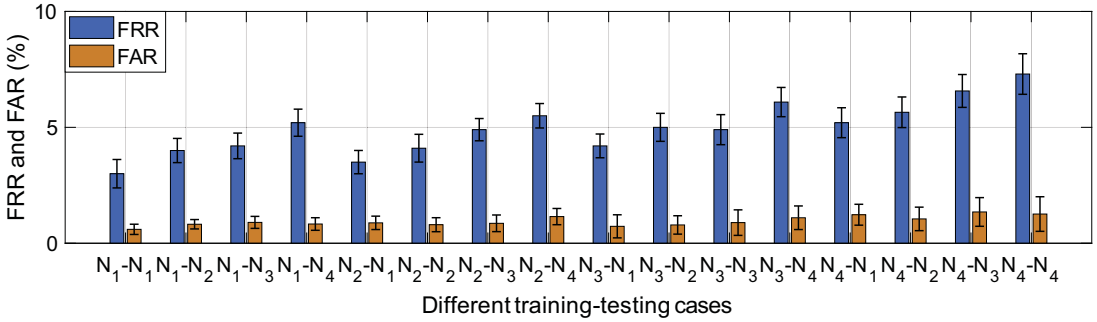
5.2 BiLock Performance

We first evaluate the overall performance of BiLock in different noise levels, and then study the effectiveness of different components and the performance across users and distances to users' lips.

5.2.1 Overall Performance. Since BiLock is an acoustic sensing based approach, it is important to evaluate its overall performance under different noise levels. Fig. 9 shows the FRRs and FARs of BiLock trained and tested under the 4 noise levels in the meeting room (*i.e.* $N_1 - N_4$, 16 *training-testing* combinations) with the tablet and smartwatch platforms. The FRRs and FARs are averaged across the 50 participants. As expected, the performance of BiLock is negatively affected by the noise levels. In the cases where BiLock is trained and tested under the same noise level, the FRR and FAR increase with the levels of noise. However, when the noise level does not exceed $50 \sim 60$ db, the FRRs and FARs almost remain constant, with average values of (4.66%, 0.88%) for tablet and (4.0%, 0.76%) for smartwatch, respectively Even when noise level reaches $60 \sim 70$ db, the performance just degrades slightly by about 3% for FRR and 0.5% for FAR. The results show the robustness of BiLock against common noise intensity. On the other hand, when the system is trained and tested under different noise levels, the average FRRs and FARs are (5.2%, 1.06%) for tablet and (5.0%, 0.97%) for smartwatch, respectively. The results are very close to those of the cases where BiLock is trained and tested in the same setting, which indicates that BiLock does not need to be re-trained even when the environment changes. Moreover, from the perspective of hardware, the average FRRs and FARs of all the 16 cases are (5.5%, 1.1%) for tablet, and (4.8%, 0.95%) for smartwatch, respectively.



(a) The overall performance of BiLock on the tablet



(b) The overall performance of BiLock on the smartwatch

Fig. 9. The overall performance of BiLock implemented on the tablet and smartwatch

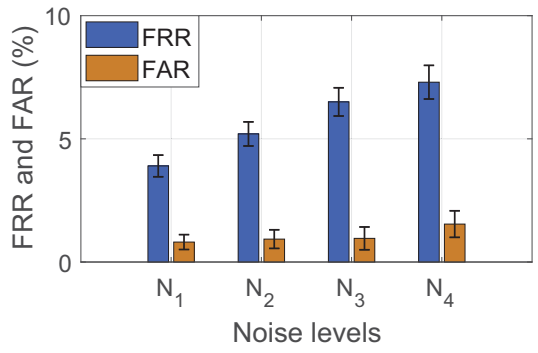


Fig. 10. Impact of mobility.

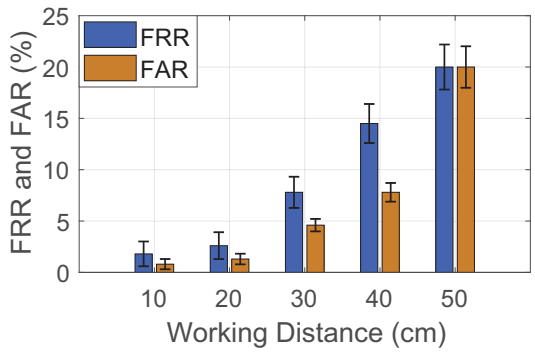
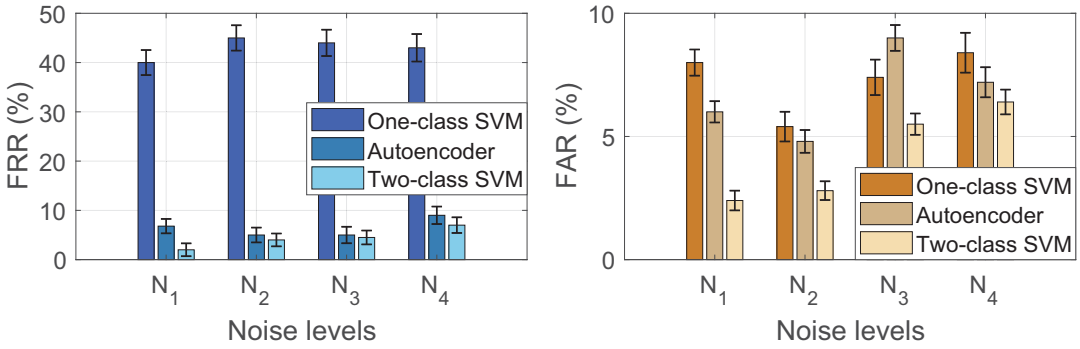


Fig. 11. Impact of distance to user's lips.

There is no notable performance gap between two hardware platforms. Therefore we only show the performance of BiLock on the smartwatch in the subsequent evaluations.

5.2.2 Performance in Mobile Scenario. Wearables are commonly used in mobile scenarios. As a result, we evaluate the performance of BiLock when it is used in motion. Similarly, we utilize data under four different noise levels to train BiLock and request participants to test it while moving around in the lab room (about 50 ~ 60 db



(a) The FRR of BiLock with different authentication models (b) The FAR of BiLock with different authentication models

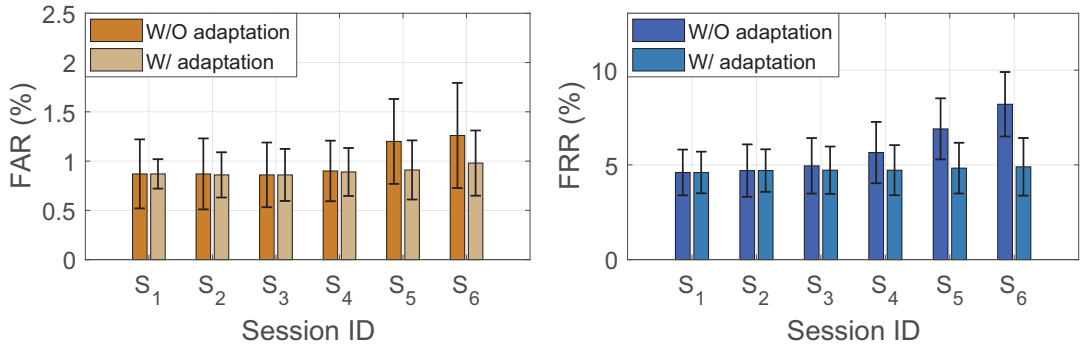
Fig. 12. Evaluation of the effectiveness of model training

noise level). Fig. 10 shows the average FAR and FRR across all users in each testing case. The average FRR and FAR for all cases are about 5.7% and 1.1%, with standard deviation of 1.3% and 0.3%, respectively. In comparison, the average FRR and FAR in cases of $N_i - N_3$ ($i = 1, 2, 3, 4$) as shown in Fig. 9(b) are 5.1% and 0.9%, respectively. The results indicate that human walking has minute impact on the performance of BiLock. Even though human walking causes some noises, they are not powerful enough to cause noticeable degradation of SNR of occlusion sounds. In our opinion, this is also an outstanding advantage over gait-based authentication systems with inertial sensors.

5.2.3 Impact of Distance to Lips. In this experiment, we place the smartwatch at different distances away from one participant's lips to test the effective operation range of BiLock. Specifically, we request each of all the participants to use BiLock with distances varying from 10 cm to 50 cm in the meeting room with 50~60 db noises. Fig. 11 shows the results. As is shown, with a larger distance to the user's lips, both the FRR and FAR increase. This is because the sounds of dental occlusion decay quickly in space. With a distance of 50 cm, the sounds are overwhelmed by background noises and BiLock fails to work. In addition, BiLock consistently achieves low FARs and FRRs when the device is placed 20 cm within the lips. Therefore, users can hold a device at any distance to their lips within 20 cm as they like without affecting system performance. It is also a reasonable operation range in practical authentication scenarios.

5.2.4 Effectiveness of Model Training. To evaluate the effectiveness of our model, we compare the performance of BiLock with different implementations. Referring to Fig. 8, we mainly consider another two feasible methods by only using a one-class SVM, and an autoencoder as authentication model. Except these changes, other settings remain the same. Fig. 12 shows the FRRs and FARs of three different approaches under different noisy environment. Comparing one-class SVM and autoencoder, the latter obviously outperforms the former in FRR and FAR. This is because autoencoder produces better representation of the input data and is more powerful in labeling the data. Moreover, we can notice that comparing with only using autoencoder as authentication model, combining it with a two-class SVM can further decrease FRR and FAR by 2.1% and 3.1%, respectively. Instead of identifying a sample by a threshold, a two-class SVM figures out a hyperplane with training data in the training stage, and then categorizes samples into either side of the hyperplane.

5.2.5 Effectiveness of Model Adaptation. To demonstrate the performance of BiLock over time and the need for model adaptation, we evaluate BiLock on 50 participants trained on the dataset of the first session (*i.e.*, S_1) and then tested in all the six sessions with 100 tests for each participant in each session. Fig. 13 shows the results



(a) The FRR of BiLock with/without model adaptation (b) The FAR of BiLock with/without model adaptation

Fig. 13. Evaluation of the effectiveness of model adaptation

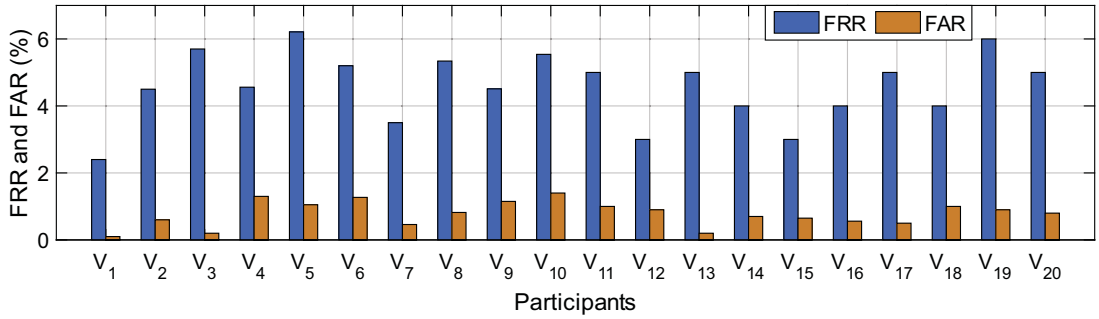


Fig. 14. Evaluation of the impact of user diversity

averaged over the 50 participants with and without model adaptation. Without model adaptation, the average FRR increases from 4.7% to 8.2%. In contrast, the average FRR remains at around 4.9% even after 3 months if model adaptation is applied. The average FARs with and without model adaptation seem insensitive to the time lags between training and testing. This indicates human dental occlusion is distinctive enough and does not change over a long time.

5.2.6 Impact of Users. We also consider the impact of user behavior diversity on system performance. Limited by page space, we only display 20 randomly selected participants' results, each of which is obtained by averaging all testing cases. As shown in Fig. 14, the maximum and minimum FRR and FAR are (6.2%, 1.4%) and (2.4%, 0.1%), respectively. This substantiates the following two main points. On one hand, different people have different dental occlusion behaviors. Someone has stronger consistence and uniqueness in dental occlusion and thus achieve lower FRR and FAR; while someone performs dental occlusion in a slightly varied way, which leads BiLock to mis-authorize users' attempts. On the other hand, the FRR and FAR do not exceed 7.0% and 1.5% even in worst cases. We notice that the performance degradation for certain participants is caused by changes of occlusion positions and intensity after long-time experiments. Multiple trials can be adopted to improve the performances of those participants.

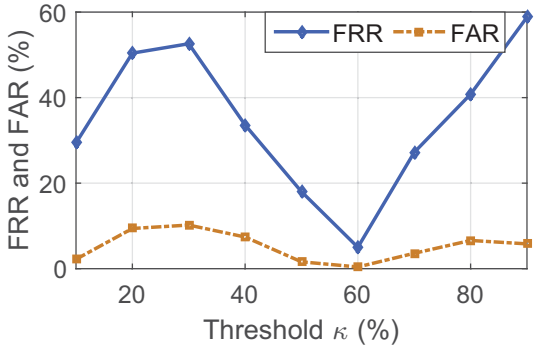


Fig. 15. Impact of threshold κ.

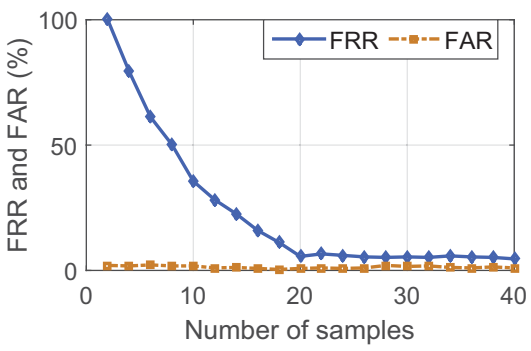


Fig. 16. Impact of # of training samples.

5.3 Impact of Parameters

5.3.1 Impact of Threshold κ. The threshold κ is an important parameter in constructing the two-class classifier during model training (see Sec. 4.4.1).In theory, a large κ filters large numbers of feasible samples and restricts data fed back for model evolution; while a small one may introduce mislabeled samples in the feedback loop, which deteriorates the accuracy of the classifier. To evaluate its impact, we vary κ in the system design and obtain BiLock's corresponding performance. Fig. 15 shows average FRR and FAR with different thresholds ranging from 0% to 100%. From the results, a medium threshold $\kappa = 60\%$ yields the lowest FRR and FAR, which is used in the aforementioned evaluations.

5.3.2 Impact of Number of Training Samples. As a user-friendly authentication mechanism, BiLock should impose little overhead to train an accurate classifier. Fig. 16 plots the FRR and FAR averaged over all the 50 participants with an increasing number of training samples. As indicated in the figure, the average FRR falls sharply before the size of training samples reaches 20. Then it stabilizes with the increase of training samples. The average FAR remains consistently low with few training samples. The results show that BiLock needs about 20 samples from a legitimate user for authentication at a reasonable accuracy. Therefore we fix the number of training samples to 20 in all the evaluations, unless specified otherwise.

5.4 Comparison with Commercial Systems

In this part, we compare BiLock with voice-based authentication methods from different aspects with three experiments. Particularly, we make use of two commercial voice-based authentication schemes which provide voice authentication services for unlocking screen or login for comparison. The first application is LockScreen which is developed on the basis of a professional voiceprint recognition engine launched by iFLYTEK, a top speech processing company [18]. The other application is WeChat's voice authentication which is used for login service. For both applications, users need to predefine their voice passwords first and speak them to the device for multiple times with microphone active. Then the applications will extract the passwords and voiceprint from audio signals and store them in the database. Only when both the password and voiceprint are matched, a person will be authenticated correctly. We first train three systems under the noise of 30 $\sim$ 40 db noise in the meeting room scenario. After that, we conduct three-part comparative experiments.

In the first experiment, we request 10 participants to perform 100 authentication trials individually with WeChat, LockScreen and BiLock under four noise levels ($N_1 \sim N_4$) in the meeting room scenario. The aim is to get knowledge of the robustness of the three systems to noise. In the second experiment, we mainly consider the security comparison of WeChat, LockScreen and BiLock against *replay attack*. In the meeting room, when

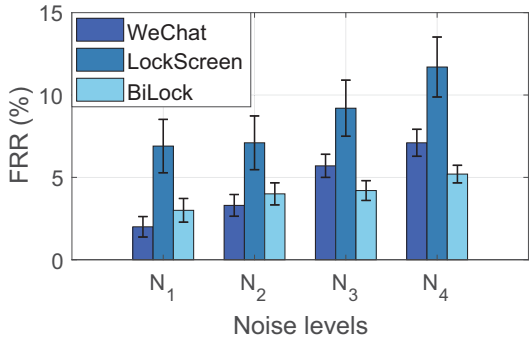


Fig. 17. The comparison of FRR between WeChat, LockScreen and BiLock under different noise levels.

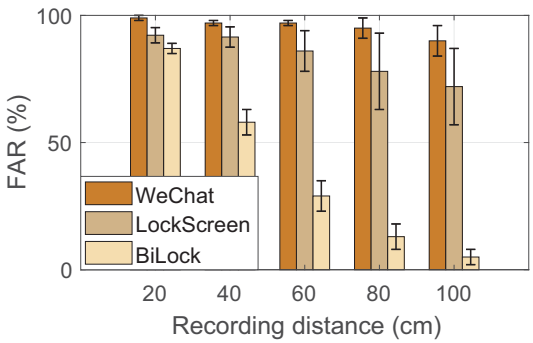


Fig. 18. The comparison of FAR between WeChat, LockScreen and BiLock at different distances.

a participant conducts 100 authentication trials with each application, we place another smartphone to record sounds with distances of 20 cm, 40 cm, 60 cm, 80 cm and 100 cm away from the site respectively. After that, we play the recorded audios to each application to pretend legitimate users. In the third experiment, we evaluate the security levels of the three methods against *observation attack*. Specifically, we randomly pick five participants to act as impostors, and another five participants (labeled as U_1 to U_5) as legitimate users. The impostors are asked to mimic the actions (*i.e.*, clicking teeth and speaking words) of the five legitimate users. For each legitimate user, each impostor attacks the "password" for 200 times per method. The impostor can observe the input of a user every 20 attack attempts, at a distance of about 0.5 m. This experiment is conducted in the meeting room scenario with 40 $\sim$ 50 db noise. All the above user studies in our work received IRB approval.

5.4.1 Robustness to External Interference. Fig. 17 shows the average FRRs of WeChat, LockScreen and BiLock over different participants under each noise level. The overall FRR averaging over all noise levels are 4.7%, 8.8% and 4.2% for WeChat, LockScreen and BiLock, respectively. As the noise level increases from 40 $\sim$ 50 db to 60 $\sim$ 70 db, WeChat and BiLock can keep a stable low FRR with a slight increase of 5.1% and 3.2% respectively, while the FRR of LockScreen increases over 7%. The results indicate that our system is comparable with WeChat and better than LockScreen in robustness to noise interference. The reason for BiLock's favorable low FRR under different levels of noise is that human voice share more similarity in frequency domain with external noise interference, and is easier to be affected by a user's emotions, health states and even speech speed.

5.4.2 Security Against Replay Attack. Fig. 18 displays the recognition results of WeChat, LockScreen and BiLock under replay attack at different distances. In such a circumstance, we regard recorded copies as impostors and thus utilize FAR as the evaluation metric. Obviously, FARs of all the three systems decrease with distance, from (99.0%, 92.2%, 87.5%) to (89.4%, 74.5%, 5.6%). With increasing distance, it is more likely for the systems to reject impostors since the SNR decreases with distance. Meanwhile, we can note that FAR of BiLock is much lower than that of WeChat and LockScreen under the same distance. The underlying reason is that sounds of dental occlusion are more close to impulse wave and decay faster than human voice in most normal cases. Due to this property, when the recording distance exceeds a certain range, it is rather difficult to differentiate occlusion sounds and background noises with a commercial microphone. As a result, we conclude that BiLock has a higher security level against replay attack compared with voiceprinting schemes.

5.4.3 Security Against Observation Attack. Fig. 19 shows the FAR of WeChat, LockScreen and BiLock under *observation attack* against five valid users. Overall, the FARs of BiLock, WeChat and LockScreen are similar, with

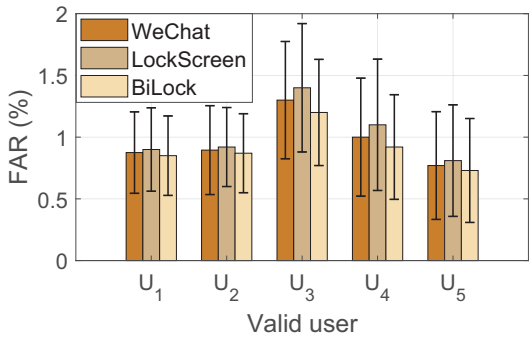


Fig. 19. The FAR of the three systems under observation attacks against five valid users.

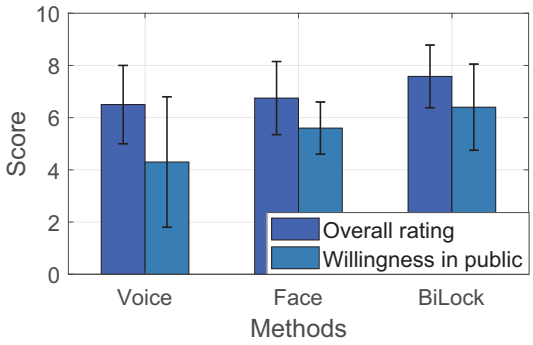


Fig. 20. The users' overall rating of different methods and willingness of using them in public scenarios.

average values lower than 1.5%. The FARs of the *observation attack* against the five valid users are consistently low (< 1.5%). Also, it is more difficult to conduct observation attack than replay attack. The reasons for these results are two-fold. On the one hand, sounds of clicking teeth mainly depend on the properties such as material, shape, mass and layout of teeth. Even the same gesture may differ in sounds across people. On the other hand, It is difficult to mimic the teeth clicking gestures by observation, because a user often keeps his/her mouth closed while performing the dental occlusion gestures.

5.5 User Survey and Feedback

In addition to validating BiLock's effectiveness, we also investigate its user experience. We conducted a survey with 100 participants. All the participants have already installed and used WeChat voice login, a popular voiceprint-based authentication scheme, and AliPay login, a popular face recognition based authentication scheme before. Among the 100 participants, 50 of them have taken part in the previous experiments (Sec. 3.2). We first informed all the participants of the aim of the study as well as the usage of BiLock. Since 50 of the participants have had experience with BiLock, they were not asked to use BiLock again. The remaining 50 were newly recruited from our campus for the user study. So they were asked to install the BiLock APK on their smartphones, and use BiLock for authentication in different daily scenarios such as libraries and lab offices. Finally, we designed a questionnaire and distributed it to all the 100 participants via a platform named Sojump, and collected feedback from each participant. The questionnaire consists of the following three questions.

- *By jointly considering the accuracy, robustness and usability, please rate the overall score of the three authentication methods (BiLock, WeChat voice login, AliPay login) from* 0 *to* 10 *(*0 *means worst;* 10 *means perfect).*
- *How are you willing to use the three methods (BiLock, WeChat voice login, AliPay login) in public e.g., library, lab office? Please rate your willingness from* 0 *to* 10 *(*0 *means I never want to use it in public;* 10 *means I would certainly love to use it in public).*
- *Can you list some advantages and disadvantages of BiLock over WeChat voice login and/or AliPay login?*

Fig. 20 shows the results of post-usage survey in terms of the overall rating and the willingness to use the method in public. The average overall ratings are 6.5, 6.8 and 7.6 for voice-based, face recognition-based and BiLock, respectively. Using the nonparametric Wilcoxon signed-rank test, the Z values and p values for BiLock and voice-based are −2.2749 and 0.0116, while those for BiLock and face recognition-based are −1.7891 and 0.03673. Hence the differences in terms of overall rating between BiLock and the other two methods are statistically

significant. The average scores of willingness to use each in public are 4.3, 5.6 and 6.4 for voice-based, face recognition-based and BiLock, respectively. Similarly, we calculate the Z-values and p-values for BiLock and voice-based authentication as -2.8563 and 0.00212, indicating statistically significant difference. However, for BiLock and face recognition based authentication, the Z-value and p-value are -1.3728 and 0.08534, meaning that the result is not significant at $p \leq 0.05$. The results show that BiLock is more acceptable than voice-based authentication in public, and is at least comparable in terms of acceptance to face recognition based method. At last, we also collect their specific comments about the advantages and disadvantages of three methods from different perspectives. We display several representative comments as follows.

- *"It is rather embarrassing to speak out words in public when using voiceprinting method. In contrast, BiLock is more imperceptible and easy to use. But I prefer to use BiLock without placing the device so near to my mouth if possible."*
- *"I use voice-prints frequently but BiLock is also cool. I think BiLock may be more robust when I caught a cold. Sometimes my phone does not recognize my voice when I got sick."*
- *"First, it is more natural than making voice especially in public places. Second, it is also helpful for the dumb and can be applied on head-mounted devices. Third, it is more secure than voiceprinting method as the intensity of tooth-click sounds is much lower than voice. However, I am not sure whether it can still keep high accuracy in more noisy environments such as crossroads, airports, concerts and* etc. *"*

6 LIMITATIONS AND FUTURE WORK

Although BiLock shows favorable properties, it has the following limitations and deserves future work towards a more practical system.

First, the robustness of BiLock under more noisy environments can be further enhanced. In the current version, BiLock can achieve high performance in environments with 60∼70 db noises which cover a wide range of daily scenarios. However, it remains uncertain whether BiLock can keep high performance in environments with more powerful noises such as busy crossroads, highways and airports. Even if its performance degrades much in such environments unfortunately, it deserves to explore more advanced denoising techniques for signal enhancement, such as spectral subtraction, adaptive filtering and the like.

Second, the implementation of model training and adaptation in BiLock can be improved. Due to the heavy burden of model training and adaptation for resource-constrained devices, current version of BiLock relies on off-line training and adaptation with model parameters determined with desktops. However, such an implementation is obviously inconvenient and cumbersome in practical scenarios. An alternative way is to upload data to the cloud, train and adapt the model, and send back model parameters to mobile devices. In this way, for devices lacking computational capability, they can afford to update the model with environmental variance. We leave this as one of our future work.

Third, the evaluation of BiLock can be more comprehensive and practical. As a concept-of-proof system, during data collection, although the participants are informed to act naturally, it is unavoidable that they perform tooth clicks a little bit mechanically. However, as indicated by [3], the system performance during live usage scenarios does not mirror that in controlled settings since tooth clicks in the two states differ slightly. As a result, from the perspective of practical use, it needs to obtain the live performance of BiLock by testing it in more practical ways in our future work. In addition, our present post-usage survey only shows a single score to assess the subjective satisfaction and acceptance in public of BiLock. We plan to design better UIs for BiLock and conduct a more comprehensive user experience study with standardized questionnaires such as UEQ [26] and AttrakDiff [16].

7 CONCLUSION

Considering shortcomings of existing methods, this paper proposes a novel biometric authentication scheme that utilizes sounds of dental occlusion as a unique feature. We demonstrate the feasibility of our scheme and also design a prototype-BiLock with embedded microphones on mobile devices to validate its effectiveness. Comprehensive experiments have shown that BiLock can achieve very low FRR and FAR even in noisy environments. In addition, compared with voiceprinting method, our scheme shows evident superiority in robustness and security. Without additional hardware, BiLock can run as a stand-alone application, or be seamlessly embedded with existing authentication system on most smart devices. We believe that the performance of BiLock can be further improved by optimizing the autoencoder and SVM models in the future work.

ACKNOWLEDGMENTS

This research was supported in part by the China NSFC Grant 61472259, Joint Key Project of the National Natural Science Foundation of China (Grant No. U1736207), Guangdong Natural Science Foundation 2017A030312008, Shenzhen Science and Technology Foundation (No. JCYJ20170302140946299, JCYJ20170412110753954), Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China(Grant No.161064), Guangdong Talent Project 2015TX01X111 and GDUPS (2015), and Tencent "Rhinoceros Birds" - Scientific Research Foundation for Young Teachers of Shenzhen University. This work is partially supported by Tianjin Key Laboratory of Advanced Networking (TANK), School of Computer Science and Technology, Tianjin University, Tianjin China, 300350. Kaishun Wu is the corresponding author.

REFERENCES

[1] Mohammed Nabih Ali, EL-Sayed A El-Dahshan, and Ashraf H Yahia. 2017. Denoising of Heart Sound Signals Using Discrete Wavelet Transform. *Circuits, Systems, and Signal Processing* 36, 11 (2017), 4482–4497.

[2] Apple Inc. 2017. iPhone X. https://www.apple.com/iphone-x/. Accessed on 2018-08-07.

[3] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jiminez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey: An exploration of tooth click gestures for hands-free user interface control. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 158–169.

[4] Debnath Bhattacharyya, Rahul Ranjan, Farkhod Alisherov, Minkyu Choi, et al. 2009. Biometric authentication: A review. *International Journal of u-and e-Service, Science and Technology* 2, 3 (2009), 13–28.

[5] Cheng Bo, Lan Zhang, Taeho Jung, Junze Han, Xiang-Yang Li, and Yu Wang. 2014. Continuous user identification via touch and movement behavioral biometrics. In *Proceedings of IEEE IPCCC*. 1–8.

[6] S Allen Broughton and Kurt Bryan. 2018. *Discrete Fourier analysis and wavelets: applications to signal and image processing*. John Wiley & Sons.

[7] Jagmohan Chauhan, Yining Hu, Suranga Seneviratne, Archan Misra, Aruna Seneviratne, and Youngki Lee. 2017. BreathPrint: Breathing acoustics-based user authentication. In *Proceedings of the ACM Mobisys*. 278–291.

[8] Adam J Cheyer. 2016. Device access using voice authentication. US Patent 9,262,612.

[9] John Chuang, Hamilton Nguyen, Charles Wang, and Benjamin Johnson. 2013. I think, therefore i am: Usability and security of authentication using brainwaves. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. Springer, 1–16.

[10] Charles K Chui. 2016. *An introduction to wavelets*. Elsevier.

[11] John Daugman. 2009. How iris recognition works. In *The essential guide to image processing*. Elsevier, 715–739.

[12] 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 ACM CHI*. ACM, 987–996.

[13] Changxing Ding and Dacheng Tao. 2016. A comprehensive survey on pose-invariant face recognition. *ACM Transactions on intelligent systems and technology* 7, 3 (2016), 37.

[14] Weixi Gu, Zheng Yang, Longfei Shangguan, Wei Sun, Kun Jin, and Yunhao Liu. 2014. Intelligent sleep stage mining service with smartphones. In *Proceedings of the ACM Ubicomp*. ACM, 649–660.

[15] Tian Hao, Guoliang Xing, and Gang Zhou. 2013. iSleep: unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM SenSys*. ACM, 4.

[16] Marc Hassenzahl and Andrew Monk. 2010. The inference of perceived usability from beauty. *Human–Computer Interaction* 25, 3 (2010), 235–260.
[17] Georg Heigold, Ignacio Moreno, Samy Bengio, and Noam Shazeer. 2016. End-to-end text-dependent speaker verification. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 5115–5119.
[18] IFLYTEK. 2017. LockScreen APP. http://www.iflytek.com/en/audioengine/list_2.html. Accessed on 2018-08-07.
[19] Anil K Jain and Ajay Kumar. 2012. Biometric recognition: an overview. In *Second generation biometrics: The ethical, legal and social context*. Springer, 49–79.
[20] Anil K Jain, Arun Ross, and Salil Prabhakar. 2004. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 1 (2004), 4–20.
[21] Thomas R Katona and George J Eckert. 2017. The mechanics of dental occlusion and disclusion. *Clinical Biomechanics* 50 (2017), 84–91.
[22] Ajay Kumar and Cyril Kwong. 2013. Towards contactless, low-cost and accurate 3D fingerprint identification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3438–3443.
[23] Ajay Kumar and Yingbo Zhou. 2012. Human identification using finger images. *IEEE Transactions on image processing* 21, 4 (2012), 2228–2244.
[24] Koichi Kuzume. 2008. A Character Input System Using Tooth-Touch Sound for Disabled People. In *International Conference on Computers for Handicapped Persons*. Springer, 1157–1160.
[25] Koichi Kuzume. 2012. Evaluation of tooth-touch sound and expiration based mouse device for disabled persons. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops*. IEEE, 387–390.
[26] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.
[27] Can Liu, Gradeigh D Clark, and Janne Lindqvist. 2017. Where Usability and Security Go Hand-in-Hand: Robust Gesture-Based Authentication for Mobile Systems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 374–386.
[28] Jian Liu, Chen Wang, Yingying Chen, and Nitesh Saxena. 2017. VibWrite: Towards Finger-input Authentication on Ubiquitous Surfaces via Physical Vibration. In *Proceedings of the SIGSAC Conference on Computer and Communications Security*. ACM, 73–87.
[29] Davide Maltoni, Dario Maio, Anil Jain, and Salil Prabhakar. 2009. *Handbook of fingerprint recognition*. Springer Science & Business Media.
[30] Judith A Markowitz. 2000. Voice biometrics. *Commun. ACM* 43, 9 (2000), 66–73.
[31] Toan Nguyen and Nasir Memon. 2017. Smartwatches Locking Methods: A Comparative Study. In *Symposium on Usable Privacy and Security*.
[32] JF Prinz. 2000. Computer aided gnathosonic analysis: distinguishing between single and multiple tooth impact sounds. *Journal of oral rehabilitation* 27, 8 (2000), 682–689.
[33] Florian Schaub, Ruben Deyhle, and Michael Weber. 2012. Password entry usability and shoulder surfing susceptibility on different smartphone platforms. In *Proceedings of the ACM Mobile and Ubiquitous Multimedia*. 13.
[34] W.J. Schull and F. Rothhammer. 1990. *The Aymara: Strategies in Human Adaptation to a Rigorous Environment*. Boston: Kluwer Academic Publishers.
[35] Muhammad Shahzad, Alex Liu, and Arjmand Samuel. 2013. Secure unlocking of mobile touch screen devices by simple gestures: you can see it but you can not do it. In *Proceedings of the Annual International Conference on Mobile Computing & Networking*. ACM, 39–50.
[36] C-S SHI and Y Mao. 1993. Elementary identification of a gnathosonic classification using an autoregressive model. *Journal of oral rehabilitation* 20, 4 (1993), 373–378.
[37] Tyler Simpson, Colin Broughton, Michel JA Gauthier, and Arthur Prochazka. 2008. Tooth-click control of a hands-free computer interface. *IEEE Transactions on Biomedical Engineering* 55, 8 (2008), 2050–2056.
[38] Tyler Simpson, Michel Gauthier, and Arthur Prochazka. 2010. Evaluation of tooth-click triggering and speech recognition in assistive technology for computer access. *Neurorehabilitation and neural repair* 24, 2 (2010), 188–194.
[39] J M Stewart. 1953. Diagnosis of Traumatic Occlusion. *The Journal of the Florida State Dental Society* (1953), 4–9.
[40] Xiao Sun, Zongqing Lu, Wenjie Hu, and Guohong Cao. 2015. SymDetector: detecting sound-related respiratory symptoms using smartphones. In *Proceedings of the ACM Ubicomp*. ACM, 97–108.
[41] Michael Tschannen, Thomas Kramer, Gian Marti, Matthias Heinzmann, and Thomas Wiatowski. 2016. Heart sound classification using deep structured features. In *Computing in Cardiology Conference*. IEEE, 565–568.
[42] David M Watt. 1969. Recording the sounds of tooth contact: a diagnostic technique for evaluation of occlusal disturbances. *International Dental Journal* (1969), 221–238.
[43] James Wayman, Anil Jain, Davide Maltoni, and Dario Maio. 2005. An introduction to biometric authentication systems. In *Biometric Systems*. Springer, 1–20.
[44] Koji Yatani and Khai N Truong. 2012. BodyScope: a wearable acoustic sensor for activity recognition. In *Proceedings of the ACM Ubicomp*. ACM, 341–350.

[45] Chen Xing Zhao, Tom Wysocki, Foteini Agrafioti, and Dimitrios Hatzinakos. 2012. Securing handheld devices and fingerprint readers with ECG biometrics. In *Proceedings of the International Conference on Biometrics: Theory, Applications and Systems*. IEEE, 150–155.
[46] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld. 2003. Face recognition: A literature survey. *ACM computing surveys* 35, 4 (2003), 399–458.

Received May 2018; revised July 2018; accepted August 2018