I Want to Know Your Hand: Authentication on Commodity Mobile Phones Based on Your Hand's Vibrations

Kevin Jiokeng, Gentian Jakllari, André-Luc Beylot

I Want to Know Your Hand: Authentication on Commodity Mobile Phones Based on Your Hand's Vibrations

KEVIN JIOKENG, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, France
GENTIAN JAKLLARI, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, France
ANDRÉ-LUC BEYLOT, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, France

We present HoldPass, the first system that can authenticate a user while they simply hold their phone. It uses the heart activity as biometric trait sensed via the hand vibrations in response to the cardiac cycle – a process known as ballistocardiography (BCG). While heart activity has been used for biometric authentication, sensing it through hand-based ballistocardiography (Hand-BCG) using standard sensors found on commodity mobile phones is an uncharted territory. Using a combination of in-depth qualitative analysis and large-scale quantitative analysis involving over 100 volunteers, we paint a detailed picture of opportunities and challenges. Authentication based on Hand-BCG is shown to be feasible but the signal is weak, uniquely prone to motion artifacts and does not land itself to the common approach of alignment-based authentication. HoldPass addresses these challenges by introducing a novel alignment-free authentication scheme that builds on asynchronous signal slicing and a data-driven algorithm for identifying a reduced set of features for characterizing a user. We implement HoldPass and evaluate it using a multi-modal approach: a large-case study involving 112 volunteers and targeted studies with a smaller set of volunteers over a period of several months. The data shows that HoldPass provides an authentication accuracy and user experience on par with or better than state-of-the-art systems with stronger requirements on hardware and/or user participation.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**; • **Security and privacy → Biometrics**.

Additional Key Words and Phrases: Authentication, Mobile Phones, Ballistocardiography, Hand

ACM Reference Format:
Kevin Jiokeng, Gentian Jakllari, and André-Luc Beylot. 2022. I Want to Know Your Hand: Authentication on Commodity Mobile Phones Based on Your Hand's Vibrations. 1, 1 (May 2022), 27 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Their name increasingly a misnomer, smartphones are becoming progressively the gateway to even the most sensitive corners of our personal and professional lives. As a result, providing access only to a legitimate user is a subject of paramount importance, as underlined by the revelations of the Pegasus Project [61]. The initial authentication systems based on knowledge factors, including passwords [46] or unlock patterns [44], are sensitive to smudge attacks and/or man-over-the-shoulder attacks [55]: it suffices for a spoofer to observe the legitimate user entering the right information for them to be able to reproduce it and access the phone. Relying on inference

Authors' addresses: Kevin Jiokeng, kevin.jiokeng@toulouse-inp.fr, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, 2 Rue Charles Camichel, Toulouse, France, 31000; Gentian Jakllari, gentian.jakllari@toulouse-inp.fr, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, 2 Rue Charles Camichel, Toulouse, France, 31000; André-Luc Beylot, andre-luc.beylot@toulouse-inp.fr, IRIT/Toulouse INP-ENSEEIHT, University of Toulouse, 2 Rue Charles Camichel, Toulouse, France, 31000.

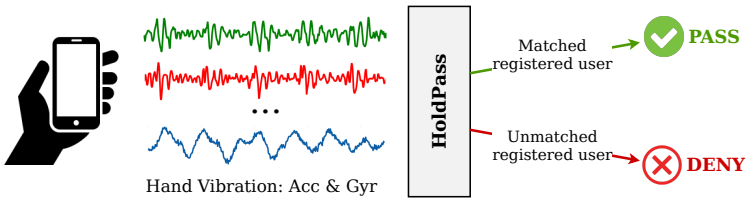


Fig. 1. High-level View of HoldPass — A novel Authentication system using natural hand vibrations (Hand-BCG signals) captured by commodity mobile phones' sensors.

Fig. 2. Experimental Setup

factors, biometrics-based systems provide a more secure authentication as they identify a user based on who they are instead of what they know. Because they offer simple and quick authentication, fingerprint scanning [52], face verification [1] and voice recognition [16] have become the most popular biometric modalities on mobile phones. The number of users relying on these modalities for authentication is rapidly increasing and expected to reach 66 % by 2024 [47]. However, these methods rely on explicit biometric modalities which are easy to observe. Recent works have shown they can be circumvented using, for example, 3D masks built with few pictures of the legitimate user [17, 34], latex fingers from the genuine user fingerprint [21, 56] or a voice recording [14]. Moreover, with the popularization of 3D printing technologies attackers can now produce more sophisticated forged biometric samples that are very difficult to detect.

To address the circumvention risk, a lot of the research emphasis recently has been placed on biometric traits that are more difficult to observe and therefore to reproduce by an attacker. Inspired by forensic dentistry, authentication on teeth alignment is proposed in [30] and occlusion sound in [70]. While marking progress in the field, smiling or making an occlusion sound may not be something people are willing to do every time they wish to unlock their phone. Leveraging vein patterns is proposed in [69] but it relies on specialized infrared sensors not found on commodity mobile phones.

Building off a rich literature showing heart activity can be used for biometric authentication [4, 12, 18, 25, 27, 37, 43], [60] proposes an authentication system based on the Seismocardiogram (SCG) signal, acquired with the phone placed on the chest. While the system is difficult to circumvent, placing the phone on the chest to unlock the phone can be awkward for some. Researchers in [41] propose a system based on the Photoplestymogram (PPG) signal, acquired with a finger placed on the phone camera. However, it does not provide enough distinctive information about the user, leading to an insufficient accuracy in realistic cross-session authentication scenarios [41].

In this work, we introduce HoldPass, a novel biometric authentication system based on the heart activity that can recognize a user while they simply hold their phone. While heart and hand seem far apart, HoldPass introduces a carefully designed architecture capable of acquiring and processing the hand vibrations – measured with the accelerometer and gyroscope sensors found on commodity mobile phones – in response to the heart activity. Figure 1 shows a high level view of HoldPass usage scenario. Fundamentally, we build on Ballistocardiography (BCG) [20], a century-old non-intrusive technique for studying heart activity based on the motion of the human body in response to the cardiac cycle. However, the BCG signal is traditionally acquired using a force sensor placed on a weighing scale or under the seat of a chair [26]. Our analysis of multiple Hand-BCG signals acquired in realistic conditions (see Sections § 2,§ 3) reveals that these signals exhibit low amplitudes, a significant level of inevitable motion artifacts and poor alignability properties, hindering the applicability of the common alignment-based authentication scheme used in systems relying on related heart activity based modalities [28, 35, 38, 60]. HoldPass therefore has to face the double challenge of identifying pertinent features that can be used for distinguishing a user that are a) alignment-free, and b) can be computed in real time on a smartphone.

HoldPass addresses these challenges by introducing a novel alignment-free authentication scheme which takes into account the particularities of the BCG signals, especially their low quality. Our design process is informed by data. We perform two large scale Hand-BCG data collection campaigns which enable us to build a large and rich dataset composed of more than 1200 measurement sessions obtained with the participation of 217 volunteers. We make this data available as an open source repository [31] for future works by the Ubicomp community. To quantify the quality of heart activity based signals for alignment-based authentication schemes and use it to analyze this large scale Hand-BCG signals database, we introduce a new metric: the Cycle Alignment Error. Our analysis reveals that common approaches for heart activity-based authentication cannot be applied to the BCG-signal. Therefore, we design an alignment-free solution involving three main steps: First, we introduce an alignment-free signal segmentation that does not depend on any event of the heartbeat cycle. Second, as Hand-BCG signals have no established fiducial points, we compute multiple feature candidates that can be used to authenticate the user and validate their usability on a large scale Hand-BCG signals database. Finally, HoldPass is faced with the challenge of reducing the feature space for running efficiently on a smartphone. A classic problem in machine learning with no general solution, we address it by flipping it on its head. First, we use validation with a classifier on a large scale database and by retro-engineering on the relation learned by the classifier we reduce the feature space through a custom feature reduction algorithm.

We implement a prototype of HoldPass and evaluate the validity of the proposed approach on our large scale open source database as well as with other targeted experiments aimed at real-world usage conditions. Our evaluation shows that HoldPass, while requiring the user to simply hold the phone, provides an authentication accuracy of 96.2 % with only 3 s of data, similar to or better than state-of-the-art heart activity based authentication systems.

To summarize, throughout this paper, we make the following contributions:

- We evaluate the feasibility of authenticating a user based on the Hand-BCG signal using a targeted qualitative analysis (Section § 2) and a large-scale quantitative analysis (Section § 3).
- We build a large scale Hand-BCG signals database consisting of more than 1200 measurement sessions including the participation of 217 volunteers, and make it publicly available for further usage by the community [31] (Section § 3).
- We introduce a new metric, the Cycle Alignment Error, to quantify the quality of heart activity based signals for alignment-based authentication schemes, and use it to analyze our large scale Hand-BCG signals database (Section § 3.3).
- We design HoldPass, a system that addresses the unique challenges raised by Hand-BCG signals in order to authenticate a user based on them (Section § 4).
- We introduce a new approach for alignment-free authentication (Section § 6).
- We implement a prototype of HoldPass and evaluate it both on our large scale open source database and other data including different authentication scenarios (Section § 7).

2 FEASIBILITY OF HAND-BCG BASED AUTHENTICATION: A QUALITATIVE ANALYSIS

2.1 Physiological sources of BCG-based authentication

Ballistocardiography (BCG) is a century-old non-intrusive technique for studying heart activity [20]. When flowing through the blood vessels, the blood ejected from the heart every systole causes the body to move in order to conserve momentum. BCG involves measuring this repetitive body motion, which can be sensed as a 3-D displacement, velocity or acceleration signal [26]. The BCG signal can be acquired on different parts of the body and using different types of sensors [26], including Inertial Measurement Units (IMU) from commodity mobile phones [36]. As an illustration, Figure 3 shows a 5 s sample of BCG signal (accelerometer z axis, gyroscope x axis) acquired with the phone in hand, as well as a reference Electrocardiogram (ECG) signal. The ECG signal is

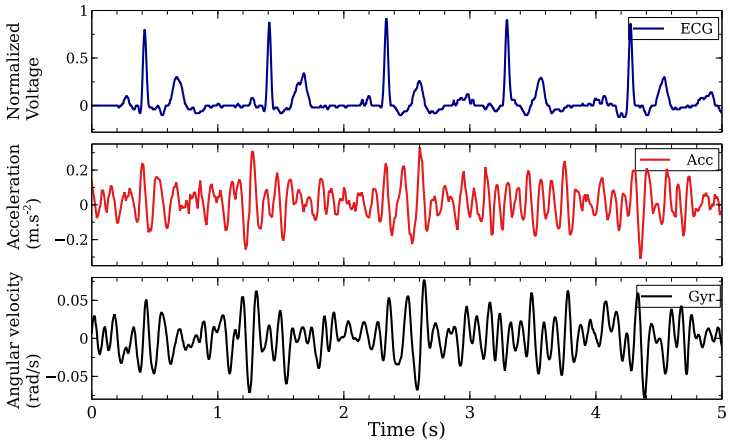


Fig. 3. BCG waveform example and ECG reference. Measured at a heart rate of 68 bpm ($\tilde{5}$.7 pulses in a 5 s window)

recorded at the same time using a GIMA PM10 Portable ECG Monitor [49]. See § 2.2 for more details about the experimental setup.

While traditionally the heart activity is studied for health purposes, recently, it is increasingly used as a biometric trait. Indeed, the heart size, shape and position differ slightly from person to person [23, 37], paving the way for authentication schemes using Electrocardiograpy [4, 27, 37], Phonocardiography [18, 25, 43] or Seismocardiography [59, 60].

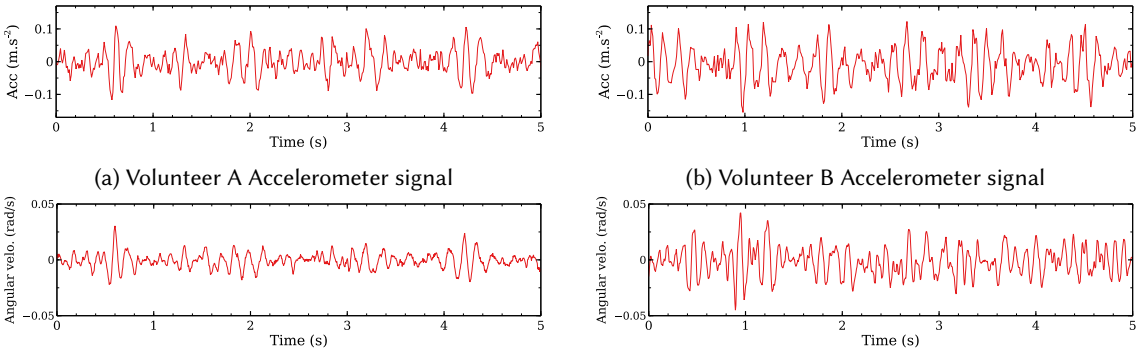
In the particular case of ballistocardiography, a study [32] into its root physiological causes revealed that the body movement is impacted by blood pressure gradients at different points in the ascending and descending aorta. Therefore, depending on the internal physiology of a person, the BCG signal will exhibit specific features reflecting their unique physical characteristics. This explains why prior studies [22, 24, 68] have been able to authenticate users using the BCG signal.

2.2 A case study

If BCG is unique for every person, it is still uncertain whether the low-quality signal measured using a smartphone in hand, an organ furthest from the heart and subject to motion artifacts, can be the basis of a reliable biometric authentication scheme. In this section, we start shedding light on this question by performing a case study with the help of two volunteers.

Experiment: Two volunteers – Volunteer A (24-year-old male) and Volunteer B (22-year-old male) – with similar heart rates (~ 70 bpm) are asked to hold a mobile phone, a Samsung Galaxy S8, while being seated with their hand placed on a table to reduce motion artifacts. The setup is shown in Figure 2. The phone runs an application recording the 3-axis accelerometer readings and 3-axis angular velocity readings from the gyroscope. The experiment lasts 30 s.

2.2.1 Time domain analysis. Figure 4 shows the first 5 s of the measured signals during the experiment. In the interest of clarity, only the axis with the highest amplitude is shown, i.e. x axis for the acceleration, and y axis for the angular velocity. The data leads to two main observations, with somehow conflicting implications regarding the feasibility of Hand-BCG-based authentication. First, Figures 4a and 4b show that despite having the same number of activity cycles, the accelerometer signals of the two volunteers exhibit quite different shapes. The signal from volunteer A presents a lower amplitude, with events better localized in time when compared to volunteer B



(a) Volunteer A Accelerometer signal

(b) Volunteer B Accelerometer signal

(c) Volunteer A Gyroscope signal

(d) Volunteer B Gyroscope signal

Fig. 4. Accelerometer and Gyroscope signals from 2 different volunteers with similar heart rates: 70 bpm (~5.8 pulses in 5 s window)

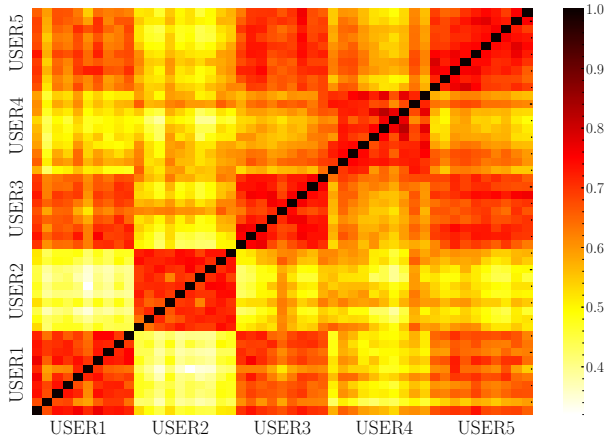


Fig. 5. Comparison of users' Hand-BCG. FFT correlation matrix.

whose signal presents more variability and higher peaks. The same observation holds for the gyroscope signals shown in Figures 4c and 4d, lending credence to the possibility of Hand-BCG-based authentication. Accelerometer and gyroscope signals provide complementary information (see evaluation in § 7.8).

However, a closer look at the signals leads to a second and more subtle observation which nevertheless can cast doubt on the basic functionality of an authentication scheme. The data shows that the different cycles from a given user are not perfectly the same and cannot be perfectly aligned, i.e. they cannot be superimposed. Some cycles have more peaks than others, and the relative heights and locations of the peaks themselves are also variable. This is more visible in the case of the gyroscope signal of Volunteer B (Figure 4d). Around t=1 s the signal presents a cycle with a significantly higher amplitude when compared to the other cycles, suggesting the existence of motion artifacts, to which the hand is uniquely prone. With authentication systems generally relying on comparing a signature on file with an input signal, the failure to align consecutive cycles of the the same signal raises significant questions of the feasibility of Hand-BCG-based authentication.

Table 1. Datasets created and used in this study. Publicly available as an online open source repository [31].

	#users	Age range	#sess per user	Session length	Phone	Acc. Fs	Gyr. Fs
Dataset A	112	20–60	10	30 s	Samsung S8	100 Hz	500 Hz
Dataset B	105	20–32	1	30 s	Google Pixel 2	50 Hz	400 Hz

2.2.2 Frequency domain analysis – Multiple users. In a second round of experiments, we extend the number of users to 5 and repeat the measurement 10 times. We transform the measured signals to the frequency domain by applying a Fast Fourier Transform (FFT) and consider all the 6 axes in order to increase the number of discriminant features. We concatenate the results from their FFT spectrums in a single vector and compare the 50 FFT vectors using Pearson correlation:

$$\rho(FFT_X, FFT_Y) = \frac{cov(FFT_X, FFT_Y)}{\sigma(FFT_X)\sigma(FFT_Y)}$$

where $cov(FFT_X, FFT_Y)$ is the covariance of the FFT vectors FFT_X and FFT_Y, and $\sigma(FFT_X)$ and $\sigma(FFT_Y)$ are their standard deviations. A higher correlation value means that two Hand-BCG signals are more similar to each other.

Figure 5 shows the obtained correlation matrix. This figure shows that data from a given user exhibit a high correlation (red squares on the diagonal), confirming, with more data, the observations made in Section § 2.2.1. At the same time, we observe that this approach can easily lead to a high False Positive Rate as signals from different users, e.g. data from USER1, USER3 and USER5 are also highly correlated. Similarly, some sessions from USER4 show a high correlation with other sessions from other users. These observations made with only 5 users show that, at a larger scale, relying only on these features would lead to a poor authentication accuracy.

2.3 Summary

Our preliminary investigation underscores a number of opportunities and challenges. The investigation of the BCG signals' physiological shows a correlation to a person's heart activity, hinting at the possibility of distinguishing users based on these signals. This is confirmed by our experimental study showing that Hand-BCG signals from each user exhibit specific features, visible both in time and frequency domains, that can be used to identify them. However, these experiments conducted with only 5 users reveal that straightforward approaches to compare these signals fail to fully capture their high complexity, casting doubt on their ability to provide the best authentication accuracy at a larger scale. In particular, the observed differences between consecutive signal cycles of the same user (section § 2.2.1) can undermine a basic functionality of an authentication system.

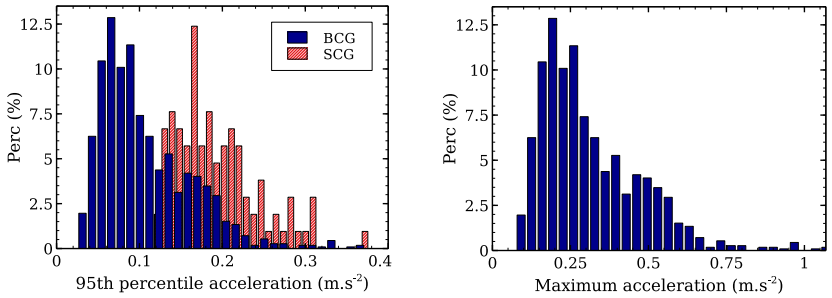
In the next section, we perform a quantitative study to explore whether these observations hold over a large population.

3 A QUANTITATIVE ANALYSIS ON THE FEASIBILITY OF HAND-BCG BASED AUTHENTICATION

3.1 Building a large-scale and open-source dataset

We performed two large-scale data collection campaigns with 217 different volunteers and involving a total of 1225 measurement sessions. All the participants are healthy students and faculty. During the data collection campaigns, each volunteer is asked to hold the phone in hand as they would naturally do, following the setup depicted in Figure 2 (Section § 2.2). The instructions are given in person by a member of our team which also verifies that the experiments are performed according to the defined protocol[1]. Each measurement session lasts 30 s during which our custom developed application records and saves sensor readings to a file containing their

[1]Our experiments are in agreement with the ethics defined in the Helsinki Declaration [15] about research involving human beings.



(a) 95th percentile peak distribution (b) Maximum peak distribution

Fig. 6. Dataset statistics (*over Dataset A*). Data collected from 112 participants, aged between 20 and 60 years old, including 19 female and 93 male subjects.

assigned anonymous identifier, e.g. USERA001. While our work relies only on accelerometer and gyroscope data, our application also records readings from other sensors that can be used for other research purposes.

The first measurement campaign is aimed at creating a dataset, Dataset A, for performing a quantitative analysis of the feasibility of Hand-BCG-based authentication as well as enabling a robust performance evaluation. As a result, we perform ten 30 s measurement sessions per user. There is a break between two consecutive measurement sessions, allowing the user to put down the phone. This is to make sure that the different measurement sessions represent as closely as possible different efforts to unlock the phone. The campaign involves 112 users aged between 20 and 60 years old, including 19 females and 93 males.

The second measurement campaign is aimed at creating a lighter dataset, Dataset B, for informing the design of our solution. As a result, we perform a single 30 s measurement session per user. The campaign involves 105 users aged between 20 and 32 years old, including 25 females and 80 males.

Table 1 summarizes the characteristics of the two datasets in terms of measurement conditions, phones, sensor characteristics and amount of data in each dataset. The data is made available for future research through an online open source repository [31]. The repository includes the raw data along with metadata providing information for facilitating their exploitation by the community.

3.2 Signal Quality Assessment

In this section, we perform an analysis of Hand-BCG signal quality based on the larger and more diverse Dataset A (section § 3.1). We start with a quality assessment based on signal amplitude and analyze how well their different cycles can be aligned or superimposed in Section § 3.3.

An analysis of the Hand-BCG signal amplitude reveals two major attributes:

Very weak signals. Figure 6a shows the 95th percentile distribution taken over the acceleration amplitude values of every measurement session in Dataset A. We consider the 95th percentile since higher values are most likely outliers due to motion artifacts. To put things into context, we add the Seismocardiogram (SCG) signals obtained with the phone on the chest from the same volunteers. The data shows that, while there is a significant variation, the signal amplitude is generally low, with an average of 0.12 m s^{-2}. It is significantly weaker compared to Seismocardiogram signals, which show a value of 0.19 m s^{-2} for the same statistics.

Motion artifacts. Figure 6b depicts the distribution of the maximum acceleration amplitude values taken over every measurement session in Dataset A. The data shows that the average value is 0.31 m s^{-2}, almost three times

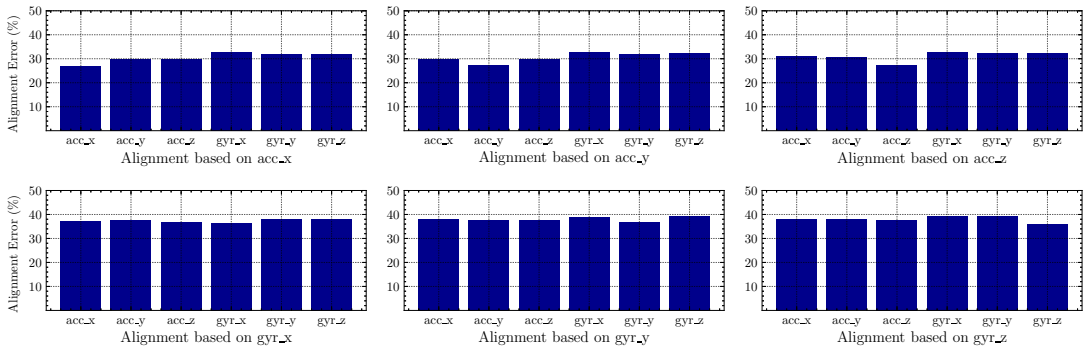


Fig. 7. Signal Alignment Error. *Because of their low quality, cycles from Hand-BCG signals cannot be reliably aligned.*

as high as that of the 95th percentile distribution, strongly suggesting these accelerometer readings are due to motion artifacts.

This observation is consistent with the intuition that holding a phone in hand inevitably leads to a high level of motion artifacts. This result, along with the weakness of the signal, highlights the challenges facing hand-BCG based authentication.

3.3 Signal Alignability

The general framework of the authentication systems involves computing a representative of the input and, at the authentication time, comparing it to the measured signal. For time-domain signals, a natural approach to select this representative is to extract a template from the signal, which is typically a cycle or a sequence of cycles [38, 60]. During authentication, the comparison is performed using alignment-sensitive metrics or features like Dynamic Time Warping (DTW) or Discrete Wavelet Transform (DWT) features. However, the case study of Section 2.2 raised doubts on whether such an approach can be applied with the hand BCG signal. In this section, we use Dataset A to perform a quantitative analysis into this question with significant implications to the design of an authentication system.

Methodology: Our alignability analysis is performed on each of the measurement sessions independently and follows a three-step procedure. First, for each signal, an axis is selected as the axis based on which the alignment is performed – the *alignment axis*. We compute the cross-correlation between the alignment axis and a fixed-length template centered at its maximum amplitude. Next, the data is split based on the peak locations in the obtained correlation signal. A minimal 0.6 s second spacing constraint is applied between the selected peaks as we consider heart rates of at most 100 bpm (beats per minutes). Finally, we evaluate the signal alignments based on the peak locations in the obtained signal slices.

Our intuition is that, for a given measurement session, a perfect alignment would lead to the maximum peaks of all the signal slices to be coherent, i.e. to appear at the same location in time – and this should happen for each of the axes, independently. To quantify the deviation from a perfect alignment, we introduce a new metric, the Cycle Alignment Error (CAE). We define the CAE of multiple signal slices along an axis as the Standard Deviation of the locations of the maximum amplitude over all the signal slices. Formally:

Definition 3.1 (Cycle Alignment Error).

$$CAE(Slices, axis) = T_s \times \underset{s \in Slices}{\text{Std}} \left\{ \underset{0 \le k < length(s)}{\text{argmax}} |s_k^{axis}| \right\},$$

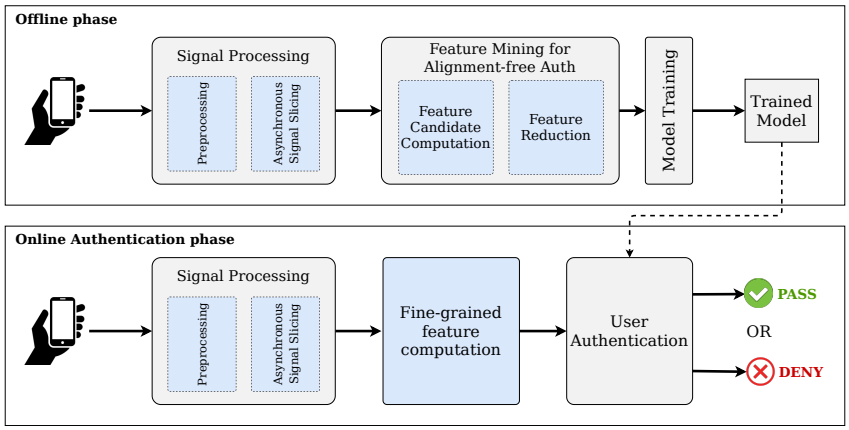


Fig. 8. HoldPass System Architecture, a Hand-BCG based authentication system

where s_k^{axis} denotes the kth value on the axis $axis$ in the signal slice s, and T_s, the signal sampling period. Our CAE metric is therefore based on correlation as we split the signal based on cross-correlation with a template before quantifying its 'alignability'. But compared to correlation which only gives 'how much a signal is present in another', our metric extends it to quantify 'how well the different cycles can be aligned or superimposed'.

We compute the CAE of all the 6 signal axes for each the measurement session in the dataset. We vary the *alignment axis* and the length of the template from 0.2 s to 1 s by strides of 50 ms, keeping the value leading to the lowest average CAE[2].

Data analysis: Figure 7 shows the normalized CAE averaged over all the measurement sessions in the dataset. The results are normalized with respect to the length of a heartbeat cycle of a subject with the median heart rate of 80 bpm (0.75 ms). The data leads to two main observations. First, we observe that the alignment error is overall extremely high, with a global average of 34.29 % (Std: 8.6 %). The second observation is that, when slicing the signal based on a given *alignment axis*, the CAE is lower along that axis but not necessarily on the other axes. Instead, the data shows that optimizing the alignment along an axis tends to have the opposite effect on the others, showing the challenge of designing a signal slicing that works for all the axes. What is more, the CAE over the *alignment axis* is still very high, with an average of 31.81 %.

To put the data of Figure 7 into context, we performed the same analysis on SCG signals collected with the phone placed on the chest. We conducted a total of 220 measurement sessions with 55 volunteers, 4 sessions by user, using a setup similar to [60]. Focusing on the accelerometer y axis, the axis of interest for SCG signals, our analysis reveals an average CAE of 12.59 %. That is less than half the value computed in Figure 7, underlining the unique challenge posed by the Hand-BCG signals.

Implication: The high alignment error revealed by our analysis makes the common alignment-based authentication approaches unsuitable for the case of Hand-BCG signal, as they are by design very sensitive to misalignment. Therefore, a new challenge emerges: designing an alignment-free authentication approach. In section § 4 we introduce HoldPass, a new authentication scheme that addresses this challenge. We compare our solution to an alignment-based authentication scheme based on DTW in § 7.4.

4 HOLDPASS SYSTEM OVERVIEW

Figure 8 shows a high-level view of HoldPass, a system that can authenticate based on Hand-BCG signals starting at any arbitrary point in time, relaxing the requirement for signal alignment. It is composed of two distinct blocks: the offline model training and the online authentication.

(1) **Offline model training:** The most involved part of HoldPass, it is aimed at training a user-specific model to be used during the online authentication stage. It includes a signal processing procedure and a novel approach for enabling alignment-free authentication. The objective of the signal processing step (Section § 5) is to a) remove all hardware specific features, and b) perform asynchronous signal slicing, paving the way for alignment-free authentication. HoldPass's approach for alignment-free authentication (Section § 6) introduces the idea that users can be distinguished based on other features that do not necessarily depend on where different events occur in time and yet still reflect the intrinsic characteristics of their Hand-BCG. As we do not know *a priori* which features can accurately capture the identity of a user, we compute multiple feature candidates in Section § 6.1 that we further reduce and fine-tune through a custom feature importance combination algorithm in Section § 6.2. With the identified pertinent features, we train a lightweight model than can accurately recognize the data of the legitimate user from the ones of a spoofer.
(2) **Online authentication:** HoldPass computes the reduced pertinent features and infer the user identity using the model trained in the previous offline phase.

In the following, we describe in detail every element of HoldPass.

5 SIGNAL PROCESSING

5.1 Preprocessing

The first step of HoldPass's processing pipeline is a preprocessing module, whose role is to remove all hardware specific features from the signal and produce a noise free version of it. It consists of two main phases: *resampling* and *denoising*.

1) Resampling: Since different phones have IMU with different sampling frequencies, the input signal – all 6 axes including acceleration and angular velocity – is resampled to a fixed sampling frequency. This allows HoldPass to run on any commodity mobile phone without modifications. To resample the signal, we use linear interpolation with a fixed time delta. It has the effect of either donwsampling, if their original sampling frequency is higher than the chosen target sampling frequency (as it is generally the case with the gyroscope readings), or upsampling the signals in the case of lower original sampling frequency. In our implementation we use $F_s = 200$ Hz as the default value of the target sampling frequency and evaluate the effect of this parameter in Section § 7.8.

2) Denoising: The aim of the denoising stage is to remove interfering signals that are added to the base BCG signal. These signals include high frequency hardware noise and the low frequency noise due to user breathing and baseline wandering. As typical respiratory rates are below 30 bpm (0.5 Hz) [2, 39] and Smartphones' IMU noise generally above 30 Hz [26, 54], we apply a fourth order Butterworth bandpass filter with cutoff frequencies of 0.5 Hz and 30 Hz. Note that this step also centers the signal by removing its mean, which is a zero-frequency component.

5.2 Asynchronous signal slicing

Faced with a signal with very high CAE (Section § 3.3), HoldPass needs to address the challenge of alignment-free authentication. The solution starts with an approach for asynchronous signal slicing. Specifically, we split the signals in slices that are independent of any specific start and end points in time. At operation time, this involves

[2]The lowest error is achieved with a template width of 0.70 s

using the data as soon as it is available, without synchronizing or aligning it to any heartbeat stage. Instead, processing starts as soon as the collected data reaches a given time length, which we denote with T_{Slice}.

Although losing some information, the asynchronous slicing has two main advantages. First, it leads to a low, constant and predictable system response time. Alignment-based schemes need to work with signal slices corresponding to specific events, for example a cycle starting at the same ATC stage of the heart cycle [60]. As a result, in practice the system is forced to wait a particular event to occur before it can start recording the input signal. In the worst case, the forced waiting time can be almost a complete cycle – 1.3 s for heart rates as low as 45 bpm. The second benefit is that it enables a better usage of the data during system development. Since the slicing no longer depends on specific event boundaries, we can split the data with a sliding window – a "cost-free" data augmentation for the subsequent processing steps. With more data, more precise and more robust relationship between the Hand-BCG signals and the corresponding user can be learned, leading to a better authentication system. We denote with T_{Slice} and T_{Stride} the length of each signal slice and the stride of the sliding window, respectively. In this study, we set the default values of these parameters to $T_{Stride} = 50$ ms and $T_{Slice} = 1.5$ s, so as to contain at least a complete cycle even for heart rates as low as 45 bpm. We evaluate the sensitivity of our approach with respect to these parameters in Section § 7.8.

6 FEATURE MINING FOR ALIGNMENT-FREE AUTHENTICATION

As the common alignment-based authentication approach is not applicable to Hand-BCG signals (see Section § 3.3), HoldPass needs to invent a new approach. It starts with the key insight that the intrinsic characteristics that can distinguish a user do not depend only on where different events appear in time, i.e. on the alignment – they can also be found in other feature domains. For instance, two users may be distinguishable because of the amplitude level of their hand vibrations, or because of how the different axes of each user's signals correlate together.

Implementing this insight however leads to two important challenges. First, we have to identify the features that can uniquely characterize a user. Second, due to the limited computational resources of commodity mobile phones, HoldPass needs to resolve the trade-off between performance and computational complexity.

One possible approach to address these challenges would be to compute features based on existing knowledge about the signal type and its physiological sources. For instance, *Cardiac Scan* [38] proposes to compute a specific number of features capturing the time differences between events of the heartbeat cycle when dealing with authentication based on rib cage reflected radio signals. This technique is widely adopted for authentication based on ECG sequences, which have well known stages and fiducial points [4, 28, 35, 48]. However, in the special case of Hand-BCG signals, there is no established prior knowledge on the shape and fiducial points of these signals. Furthermore, these approaches are based on intuition, without proof of whether the selected features are the best for authentication.

HoldPass adopts a data-driven and progressive approach. We start by computing multiple feature candidates and train a classifier to predict the user based on the computed features (Section § 6.1). We build on this to address the accuracy-computational complexity tradeoff by leveraging the relationship learned by the classifier and applying careful retro-engineering (Section § 6.2).

6.1 Learning from Multiple Features

6.1.1 Feature Candidates. To boost the odds of identifying a maximum number of pertinent features, we compute multiple features from different representation domains:

- **Spectral features:** We compute the Fourier transform of the signal and keep the resulting coefficients for all the frequencies. We also aggregate them by taking their mean, median, variance and standard deviation. To have another view of these spectral components, we also compute the entropy of the Power Spectral Density and its binned entropy.

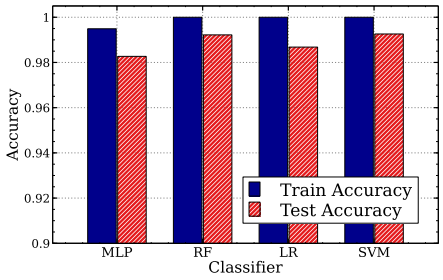


Fig. 9. Authentication Accuracy with All Features

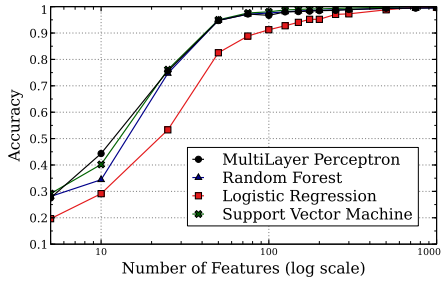


Fig. 10. Authentication Accuracy with Reduced Features

- **Moment and Distribution features:** This includes the median, variance, standard deviation, skewness, kurtosis, minimum, maximum, root mean square and different quantiles. We also include the number and percentage of values that are above or below given thresholds (mean, median and static thresholds), and the ratio between the standard deviation and the range of observed values (max - min).
- **Energy and Complexity features:** We compute the maximum amplitude of the signal, its total energy, its sum of changes and its absolute sum of changes. We also add the number of peaks and valleys in the signal and the prominence of each of them.
- **Predictiveness and Trend features:** This includes the approximate entropy of the signal, its binned entropy, sample entropy and permutation entropy. We augment these features with the pvalue, rvalue, intercept, slope and standard error of a fitted linear regression model; and also the Benford correlation of the signal and its c3 statistics value.
- **Self-axis variation features:** This includes the auto-correlation coefficients of each axis with itself and their aggregated statistics: mean, variance, standard deviation, median. We also compute the auto regressive coefficients and the first and second order derivative of the signal.
- **Cross-axes variation features:** Different from the other features that are computed for each axis independently[3], we also compute features that captures how the different axes vary with respect to each other. For this purpose we compute the cross-correlation between each pair of the accelerometer signals on the one hand, and of the gyroscope signals on the other hand.

Note that for computations that report a list of values, e.g. Fourier transform, we consider each element in the list as a different feature and we pack all of them alongside the other features to form the extracted feature vector. As a result, a total of 5,322 features are computed for each signal slice.

6.1.2 General Model Training.

Methodology. Our first goal is to observe whether different users can be distinguished based on the candidate features computed in the previous section. To this end, we train different classifiers whose aim is to predict the user with these features as input. We use Dataset B including 105 users (Section § 3.1) and preprocess and slice the data as described in Sections § 5. With 29 s of data[4] we get 551 signal slices per user. After computing the features described above, we get a total dataset size of 57,855 samples × 5,322 features. We randomly split the data into two subsets, with 70 % and 30 % of the data for the training and test sets, respectively.

We train and compare the results given by four different classifiers: a Random Forest classifier (RF), a Multi-Layer Perceptron classifier (MLP), a Logistic Regression classifier (LR) and a Support Vector Machine classifier

[3]See https://tsfresh.readthedocs.io/en/latest/text/list_of_features.html for an exhaustive list of them [13]
[4]We exclude the first second during which the measured signal is sometimes abnormally high because of the user's hand being not yet stable enough

(SVM). Before fitting the classifiers to the data, we scale each feature by subtracting its mean and dividing with its standard deviation, accelerating the training and improving the results, To find the models that lead to the best performance, we vary the parameter values of each of those classifiers and apply a five-fold cross-validation with a standard grid search procedure. This procedure also serves to combat overfitting. The varied parameters are:

- the number of trees, maximum depth, number of features considered for each split and the criterion for the RF classifier;
- the number of layers and number of neurons per layer along with the activation and the regularization parameter for the MLP;
- the regularization parameter for the LR;
- the regularization parameter and the used kernel for the SVM classifier.

At the end of this cross-validation procedure, we keep the model which leads to the best performance on the training set, and evaluate it on the test set.

Data: Figure 9 shows the results obtained with the different classifiers in terms of accuracy on the training and test sets. Independently of the classifier in use, the users are almost perfectly identified based on the computed features – with an accuracy approaching 100 % both on the train and test sets. The minimum observed accuracy is 98.27 % (obtained by the MLP classifier on the test set). Similarly, the *F-Score* is always very high (minimum of 98.2 %), showing a good authentication capability on this large dataset.

Takeway: The data suggests that users can be accurately distinguished based on the computed features when using the Hand-BCG signals. Nevertheless, as the number of computed features is high – 5,322 for each signal slice of 1.5 s – the time taken to compute them is also high. A multi-threaded implementation of this computation takes on average 0.83 s on a recent Dell Latitude 5480 computer with a 16-thread CPU and 16 GB of memory. Using all of those features would lead to a very high response time of the system, especially on commodity mobile phones with their limited computing capabilities.

In Section § 6.2, we introduce how HoldPass reduces the computational complexity by reducing the feature space while still maintaining a good accuracy.

6.2 Feature Reduction

Reducing the dimensions of features without sacrificing authentication performance serves two main purposes. First, it enables a fast feature computation at the operation time and leads to a less complex model that can run rapidly on an off-the-shelf mobile phone. Second, it improves the generalization capability of the system which can be trained on a reasonable amount of data, avoiding the curse of dimensionality issue [57].

To address the computational complexity-accuracy tradeoff we introduce an approach that leverages the relationship learned by the classifier in the previous step (§ 6.1) and retro-engineers it to find the features having the greatest impact on the decision. Specifically, we perform a model inspection and study the different parameters learned by the classifier to make its predictions.

6.2.1 Grouped Feature Importance. While it can be applied to all the trained models, we perform the model inspection only on the Random Forest classifier (RF). The reason being that its easy-to-understand internal functioning lands itself to better retro-engineering. We evaluate the other models on the features selected with RF in Section § 6.2.2.

Our approach using the RF classifier, summarized in Algorithm 1, is as follows. In the first step (lines 4-6), we compute the importance of each individual feature in the candidate feature set. Specifically, for each feature in the feature space, we compute its importance as the Mean Decrease Impurity (MDI) [6] caused by this feature during the training of the RF model. Let N_f^t denote the set of node splits in tree t that include feature f, $M(n)$ the number of samples reaching node n, M is the total number of samples in the training set and T the set of trees in the RF. Then:

Algorithm 1: Feature Reduction

1 **ReduceFeatures** *(Features, RF)*
  Input: *Features*: A set of feature candidates
    RF: A trained RF model using *Features*
  Output: *FD*: Subset of most relevant feature domains
2 Set T to the set of trees in RF
3 Set M to the size of the training set
4 **foreach** $f \in Features$ **do**
5 Compute importance of f Using Equation 1;
6 **end**
7 Group features by domains: $FD = group(Features)$;
8 **foreach** $d \in FD$ **do**
9 Compute the grouped importance of d as $GFI(d) = \sum_{f \in d} MDI(f)$;
10 **end**
11 Sort FD in descending order of GFI: $FD = sort_{GFI}(FD)$;
12 Select the K most important domains that lead to a good accuracy: $FD = FD[1:K]$;
13 **return** FD;

$$MDI(f) = \frac{1}{M}\frac{1}{|T|}\sum_{t \in T}\sum_{n \in N_f^t} M(n) \times DI(n) \tag{1}$$

where $DI(n)$ is the decrease in impurity when splitting data at node n, computed as the difference between the impurity of data reaching node n and the sum of impurities at its children nodes[5].

MDI quantifies the intuition that the features which split the data more efficiently (in terms of impurity) are more crucial to the classification decision than the others. Further, it takes into account the fact that the features intervening at the lowest depth, closer to the tree roots, or which are used at multiple nodes and in multiple different trees are more important than the others.

In the second step (lines 7-10), the computed features are aggregated by domains. In a key decision driven by the goal of striking a good accuracy-computational complexity tradeoff, the algorithm shifts to identifying the best feature domains instead of individual features. For instance, when considering the Fourier coefficients domain, even if only one half plays a very important role in the classification decision, the algorithm keeps the entire set. With such computations being atomic, that is, returning all the features in a row, keeping the least important half does not impact the computational complexity while nevertheless contributing to the accuracy. We approximate the grouped importance, $GFI(d)$, of a feature domain, d, as the sum of the importance of all the features in the domain,

$$GFI(d) = \sum_{f \in d} MDI(f)$$

Finally, the feature domains are sort in descending order of their contribution and, by cross-validation on the training set, we select the minimal subset of them that leads to the best accuracy (lines 11-12).

[5]We use the Gini impurity [53] as the impurity measure.

Table 2. System parameters and their default values

Parameter	Range	Default Value
Sampling frequency	100 - 1000 Hz	200 Hz
Length of a signal slice	0.5 – 5 s	1.5 s
Stride of the sliding window	50 ms – 1.5 s	50 ms
Number of benchmark users	1 – 50	30
Number of registration sessions	1 – 5	5
Length of a registration session	3 – 30 s	30 s
Sensors in use	Acc, Gyr, Acc + Gyr	Acc + Gyr

6.2.2 Results with Reduced Feature Subset. To illustrate the effectiveness of the feature selection approach, we train and test different classifiers using the output from our algorithm. Specifically, we select the most pertinent features and train different classifiers using only them. We use the same classifiers and apply the same procedure as in section § 6.1.2, while varying the number of the top features selected from the candidate features. Figure 10 shows the accuracy on the test set as a function of the number of features in the selected feature subset. The data shows that, apart from the LR classifier, all the classifiers reach a very good identification accuracy with less than 100 features. The LR classifier crosses the bar of 95 % accuracy with 400 features, which is still very low – two orders of magnitude lower – compared to the 5,322 features in the original dataset.

This result shows that, regardless of the RF used to identify them, the selected features lead to a good accuracy for all classifiers, showing that they capture the unique characteristics of each user's Hand-BCG. In the next section, we perform a thorough evaluation.

7 EVALUATION

7.1 Implementation

We implemented HoldPass as a standalone Android application running on commodity mobile phones. Our prototype records the accelerometer and gyroscope Hand-BCG signals and processes them according to the steps depicted in Fig. 8 to decide whether the user holding the phone is legitimate or a spoofer. The signal processing and feature computation parts are implemented in standard Java and make use of *Apache Commons Maths* library [3] and the Logistic Regression is implemented with *Weka* Machine Learning library [62].

7.2 Training and evaluation methodology

To evaluate the performance of HoldPass, we use the data from the large and diverse Dataset A (Section § 3.1) comprising 1120 measurement sessions with 112 different users. During the evaluation process, we place ourselves in a realistic scenario where we train a classifier for each user in the dataset and evaluate how well HoldPass is able to recognize that user among attempting spoofers. We train a Logistic Regression classifier for each user and set its regularization strength through a five-fold cross-validation procedure on the training set. We carefully split the data by users and sessions as described below in such a way to avoid overfitting and perform a realistic evaluation. The evaluation is performed offline on a computer.

User split. We apply a leave-k-users-out strategy, training the user model against *n_benchmark_users* randomly selected users, and evaluate the trained model on the remaining ones. Thus, each user acts once as the legitimate user and is evaluated against 81 spoofers that have never been seen by the system (we set the default value of *n_benchmark_users* to 30). In the Logistic Regression training process, the samples from the legitimate user serve as the positive samples while those from the benchmark users as the negative. To ensure a fair evaluation, we randomly select as many spoofer samples as there are legitimate user samples in the test set.

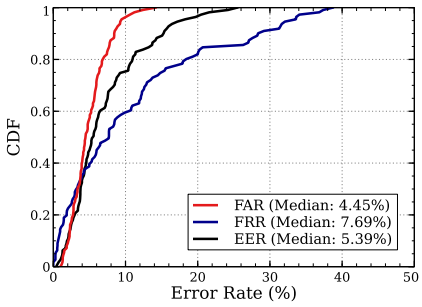


Fig. 11. Overall Performance of HoldPass

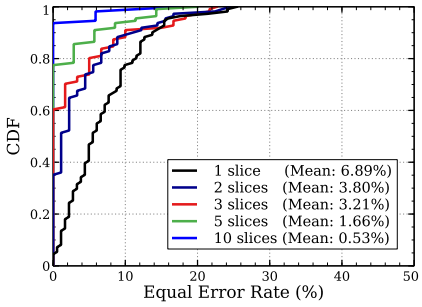


Fig. 12. Time is security: Using multiple signal slices improves significantly HoldPass's authentication performance.

Legitimate user data split. For the legitimate user data, we perform a sessions-wise split in such a way that the evaluation is performed on data coming from measurement sessions that have not been seen during training. This is to avoid a biased evaluation and to emulate the production scenario wherein the system is trained once with some pre-registered data and used later with (unseen) live data points. By default, we set the number of the registration sessions to 5, corresponding to half the user data.

Advanced datasets. In a second round of experiments, we select two groups of volunteers that were part of the first data collection campaigns (Dataset A) to evaluate the robustness of HoldPass. With the first group composed of 12 volunteers, we perform different evaluations related to usage scenarios, including assessing the authentication performance after a long period of time (Section § 7.7) and testing under different experimental conditions (Section § 7.9). The second group is composed of 20 volunteers and serves an advanced evaluation with more diverse and challenging experimental setup (Section § 7.9).

System configuration. Table 2 summarizes the system parameters that we use in this evaluation. If not explicitly specified, we set these parameters to their default values. We evaluate the sensitivity of the system to these parameters in section § 7.8.

7.3 Evaluation Metrics

Throughout the evaluation process, we employ the following performance metrics built around the decisions of Accept (True) and Reject (False). An authentication effort prompts an Accept (True) decision by HoldPass if the probability of it being from the legitimate user is above the *acceptance threshold*, Reject (False) otherwise.

- **False Accept Rate (FAR) = FA/(FA+TR)** – the probability that the system incorrectly authenticates a spoofer as a legitimate user.
- **False Reject Rate (FRR) = FR/(FR+TA)** – the probability that the system incorrectly authenticates the legitimate user as a spoofer.
- **Equal Error Rate (EER):** By varying the acceptance threshold, one can improve the FAR at the expense of the FRR, and vice versa. EER is the point where FAR is equal to FRR.
- **Accuracy (Acc.)**: the probability that the system correctly identifies both the legitimate user and the spoofers. If not otherwise specified, we compute the accuracy at the Equal Error Rate point, i.e. $Accuracy = 1 - EER$.

7.4 Overall Performance

We start with the evaluation of HoldPass' overall performance over the entire dataset while setting the acceptance threshold to its default value of 0.5. Figure 11 shows that HoldPass is able to accurately recognize the legitimate

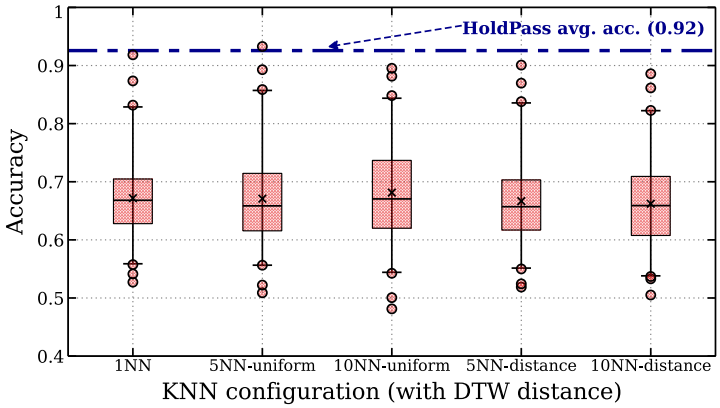


Fig. 13. Comparison of HoldPass against alignment-based authentication scheme (KNN with DTW distance)

user among spoofers, achieving a median FAR of 4.45 % and median EER of 5.39 %. In addition, for more than 90 % of the participants the FAR and EER are less than 8.5 % and 13.75 %, respectively. On average, HoldPass achieves an accuracy of 92.57 % (EER of 7.43 %).

In addition to using a single asynchronous signal slice for making a decision, HoldPass can combine the likelihood of multiple consecutive signal slices. In this configuration, the system accepts the attempting user as legitimate if the average likelihood for a certain number of consecutive signal slices is above the acceptance threshold. Figure 12 shows the CDF of the EER for different values of signal slices. For each user and each number of signal slices, we vary the acceptance threshold and adopt the value leading to the EER. The data shows that HoldPass's performance can improve dramatically when leveraging multiple signal slices. The accuracy improves to 96.2 % with only 2 consecutive signal slices (3 s of data) and reaches 97.95 % and 98.34 % when using 4 and 5 of them, respectively. In practical terms, this results provides users a virtual knob for setting the desired security-response time trade-off.

To put this performance of HoldPass into context, we performed a comparison against an alignment-based method in the time domain. For this purpose, we implemented a standard k-nearest neighbor (KNN) classifier based on state-of-the-art Dynamic Time Warping (DTW) distance. We build the dataset by splitting the signals in cycles based on cross-correlation. We vary the number k of neighbors participating in the decision ($k \in \{1, 5, 10\}$) and also vary the weighting function between *uniform* (the class votes of the selected neighbors are considered with the same weight) and *distance* (the class votes are weighted by the inverse of their distance to the input sample). We assess the accuracy achieved by the KNN classifier in each of the 5 possible configurations and for each of the users. Figure 13 shows a boxplot of the obtained accuracy, for each of the configuration. The whisker mode of the boxplots is 2/98 percentile. The data shows that, in an important fraction of the cases, DTW fails to accurately discriminate between the legitimate user and an attacker, yielding to an accuracy always under the 70 % bar. The best result, achieved with uniform votes from the 10 nearest neighbors, is an accuracy of 68.13 %. With its 92.57 % accuracy in single slice configuration, HoldPass outperforms this pure time domain authentication scheme.

7.5 HoldPass User Experience

In this section, we evaluate the user experience when using HoldPass as a smartphone user authentication system. We evaluate this experience in terms of the number of attempts for a successful login, amount of registration data needed for a good performance and system response time.

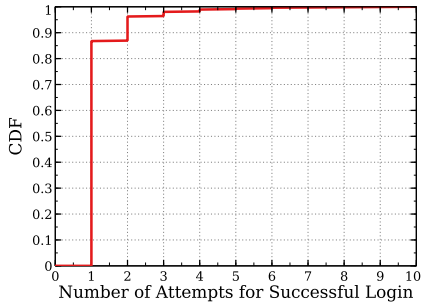


Fig. 14. CDF of number of attempts for successful login.

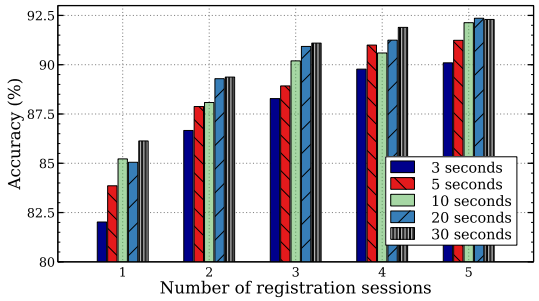


Fig. 15. Accuracy with different amount of registration data.

Table 3. Response time of the system

Phone	Signal processing	Feature computation	Logistic Regression	Total
Samsung Galaxy S8	3.8 ms	27.4 ms	4.5 ms	35.7 ms
Google Pixel 4a	3.1 ms	14.2 ms	2.9 ms	20.2 ms
OnePlus 8T	4.3 ms	28.9 ms	4.9 ms	38.1 ms
Average	**3.7** ms	**23.5** ms	**4.1** ms	**31.3** ms

Number of attempts for successful login. We evaluate the number of attempts required for a successful login for each of the 560 testing sessions in the dataset (5 for each user). For every user, we set the acceptance threshold to the one that leads to the EER and show the resulting CDF over all the users in Figure 14. The data shows that more than 85 % of the login operations are successful after a single attempt. After 2 and 3 login attempts, the user is able to unlock their phone in more than 96 % and 98 % of cases, respectively.

Amount of registered data for training. We evaluate the amount of data that need to be registered for HoldPass to have a good performance. To this end, we vary the number of registration sessions used for training the system from 1 to 5 and the length of each registration session from 3 to 30 seconds. Figure 15 shows that HoldPass' performance is an almost monotonically increasing function of the two parameters: the number of registration sessions and length of each session. However, we observe that increasing the number of registration sessions has a greater impact on the performance of the system as it enables HoldPass to better capture the variability of the Hand-BCG signals. Most important, the data shows that HoldPass is able to cross the bar of 90 % accuracy with only 15 s of registration data (5 sessions × 3 seconds) – a reasonable one-off requirement for using HoldPass.

Response time. We evaluate the time taken by HoldPass to perform the authentication on three commodity mobile phones including a Samsung Galaxy S8 (2017), a Google Pixel 4a (2020) and a OnePlus 8T (2020) and report the results in Table 3. The data shows that HoldPass runs in real time, with an average processing time as low as 31.3 ms. Added to the time taken to record the signal, HoldPass achieves a reasonable average response time of 1.53 s. As a reference, the official face identification on the recent iPhone X takes approximately 1.5 s to authenticate a user [50]. In section § 7.8 we show that HoldPass still works well with signal lengths as short as 0.5 s, therefore providing a better user experience.

Table 4. Comparison of HoldPass with other Heart Activity based Authentication systems. *HB: Heartbeats. NS: Not Specified. TAR: True Accept Rate. FAR: False Accept Rate.*

System	Device	Ref. signal	#Users	Signal length	Accuracy
System in [24]	Smart Eyewear	Head-BCG	12	3 s	96.5 %
System in [68]	Custom BCG chair	Body-BCG	91	10 HB (~8 s)	99.1 %
System in [22]	Custom BCG chair	Body-BCG	25	10 s	96 %
System in [60]	Mobile Phone	Chest-SCG	20	5 HB (~4 s)	96.49 %
System in [41]	Mobile Phone	Finger-PPG	15	NS	92 %
System in [4]	Mobile Phone + ECG sensor	Finger-ECG	10	4 s	NS. TAR: 81.82 %, FAR: 1.41 %
HoldPass	**Mobile Phone**	**Hand-BCG**	**112**	**3 s (2 slices)**	**96.2 %**

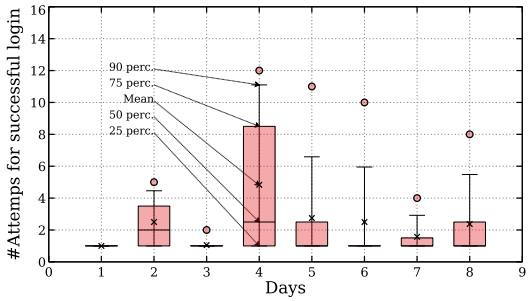


Fig. 16. Number of attempts for successful login over 8 working days.

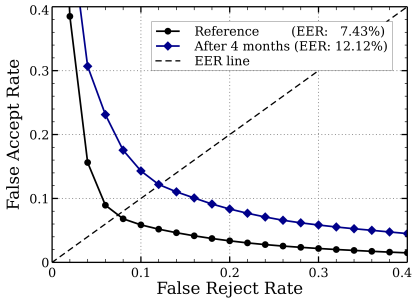


Fig. 17. Performance four months later

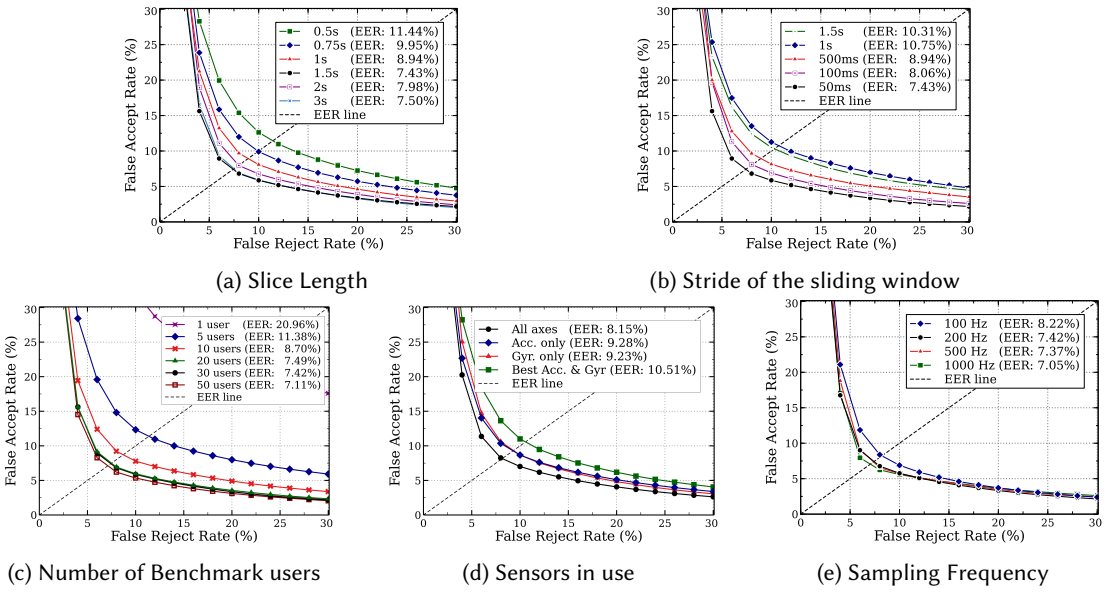
7.6 Comparison of HoldPass with other Heart Activity based Authentication Systems

In this section, we aim at contextualizing the performance of HoldPass by comparing it with other heart activity based authentication systems. To this end, we compare HoldPass to authentication systems using BCG as reference signal [22, 24, 68] and to other authentication systems based on heart activity measured with mobile phones. The latter category includes the system in [60] which uses the Seismocardiogram (SCG) signal acquired with the phone placed on the chest, the system in [41] which uses Photoplestymogram (PPG) signal acquired with the user placing their finger on the phone camera and the system in [4] which relies on Electrocardiogram (ECG) signal measured with a ECG sensor added to the mobile phone.

Table 4 shows that HoldPass achieves an accuracy comparable to the state of the art while requiring less or equivalent amount of signal data. Moreover, HoldPass does not require any custom hardware and is very convenient to use as the user just needs to hold their phone to be authenticated.

7.7 Performance Over Time

In this section, we evaluate the performance of HoldPass over time in order to assess the permanence of a user's Hand-BCG. For this purpose, we select 12 volunteers who were part of the large dataset collection campaign (Dataset A) and first perform measurement sessions over 8 consecutive working days. Four months later, for each volunteer we perform data collection and authentication attempts with the system configured to their previously identified EER threshold. This process is done offline on a computer.



(a) Slice Length (b) Stride of the sliding window

(c) Number of Benchmark users (d) Sensors in use (e) Sampling Frequency

Fig. 18. Tradeoff Detection Error. Effect of different system parameters

Figure 16 shows the distribution of the number of login attempts after which the volunteer was recognized as the legitimate user over 8 days. The data shows that a user can unlock their phone after a few attempts, with only 1.96 attempts on average and a 90^{th}percentile of 4.

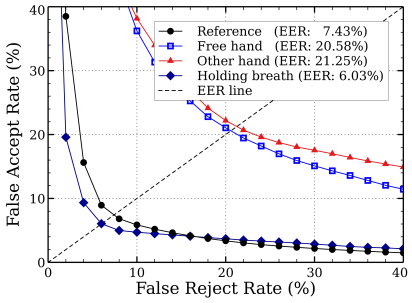
Figure 17 shows the detection error when evaluating HoldPass on the data measured four months later. We observe that even after a significant period of time HoldPass still maintains an 87.88 % accuracy. This can be improved further by performing continuous or periodic model updates based on data from successful logins.

7.8 Sensitivity Analysis: Effect of system parameters

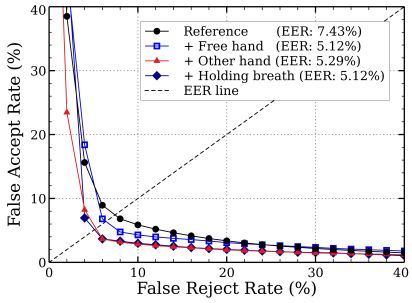
In this section, we evaluate the sensitivity of HoldPass to different system parameters. To this end, we set all the parameters to their default values, as shown in Table 2, and re-execute the main evaluation scenario (section § 7.1) while varying the value of each parameter independently. Figure 18 shows the tradeoff detection error curves. The default HoldPass configuration are plotted in black with circle markers. All the results are obtained using a single signal slice to perform the authentication.

Effect of the slice length (T_{Slice}). Figure 18a shows the results obtained while varying the signal slice length. This figure shows that the error decreases with the length of the slice length. With 0.75 s of data, HoldPass achieves an accuracy above the bar of 90 % and reaches 92.5 % with 3 s of data. Comparing to the results in the second paragraph of fig. 11 (96.2 % with 1.5 s × 2 slices), we observe that the performance of HoldPass is higher when performing the decision based on multiple signal slices.

Effect of the stride of the sliding window (T_{Stride}). Figure 18b shows the results obtained with different values of the stride applied when splitting the signals in sliding windows. If not applying a sliding window (stride 1.5 s, equal to the signal length), the accuracy of HoldPass is 89.69 %. On the contrary, we observe that lower values of this parameter lead to a better performance of the system. This result demonstrates how the alignment-free authentication scheme introduced by HoldPass can take advantage of a better utilization of the data.



(a) Different experimental conditions



(b) Trained with data from different experimental conditions

Fig. 19. Performance in different experimental conditions

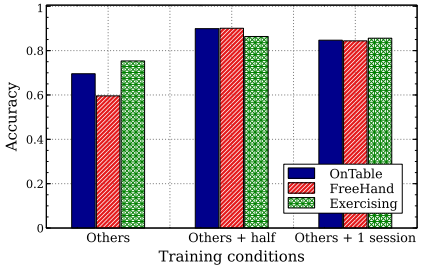


Fig. 20. Cross conditions testing

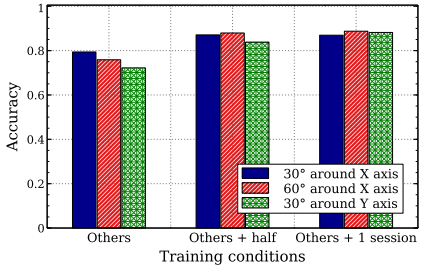


Fig. 21. Holding the phone with different orientations

Effect of the number of benchmark users. We evaluated the performance achieved by HoldPass as function of the number of benchmark users against which each legitimate user model is trained. From Figure 18c, we observe that the error decreases when HoldPass is trained against data from an increasing number of users but almost stops improving for higher values of this parameter (7.11 % for 50 benchmark users). With more benchmark users, and thus more varied spoofers data, the system is able to learn a better decision boundary to identify the legitimate user's Hand-BCG features. The performance yet remains reasonable even with lower values of this parameter, highlighting a good scalability of the system. HoldPass achieves an accuracy as high as 91.30 % when trained against 10 benchmark users only.

Effect of the sensors used by the system. By default, HoldPass uses both the acceleration and the angular velocity measured when the user holds their phone. Figure 18d shows the results obtained when relying on the accelerometer or gyroscope only. We also compare with the configuration where HoldPass would use the best axis of each of these sensors, which are selected as the ones with the highest amplitude. The best result is achieved when using the data from all the axes. When using only the accelerometer or only the gyroscope, HoldPass achieves an accuracy of 90.62 % and 90.67 %, respectively.

Effect of the sampling frequency (F_S). Figure 18e shows that the sampling frequency has very little impact on the performance of the system. HoldPass achieves an accuracy of 92.95 % and 91.78 % when the value of this parameter is set to 1000 Hz and 100 Hz, respectively.

7.9 Performance in different experimental conditions

In this section, we evaluate the performance of HoldPass in different experimental conditions, including holding the phone in the non dominant hand, holding breath and without any support under the hand. For this evaluation, we perform new measurements with the help of 12 volunteers. Figure 19a shows that breathing has a little impact on HoldPass' performance, with the accuracy slightly improving when the user holds their breath. On the other hand, the data shows that the accuracy drops when changing hand or holding the phone without any support. This can be explained by the fact that the data measured in these scenarios are quite different from the reference ones and have not been included in the system training. To corroborate, Figure 19b shows that this can be corrected by including signals from these experimental conditions in the registration data. In these configurations, HoldPass achieves a similar and performance when compared to the reference single-condition version.

To further extend the evaluation of HoldPass, we repeated the same protocol with 20 volunteers performing measurement sessions in different experimental conditions : i) with the hand on a supporting table, ii) free hand without any support and iii) after physical activity (on supporting table). In the following, we refer to these conditions as *OnTable*, *FreeHand* and *Exercising*, respectively. Here, we aim at testing how HoldPass performs when tested on data from an unseen condition. For this purpose, we perform cross conditions testing where, for each user and each target condition c_t, we train HoldPass with data from other conditions $c \neq c_t$ and test it on c_t, eg. training on data from *OnTable + FreeHand* and testing on *Exercising*. We also evaluate the case where HoldPass is trained with data from all conditions, either with half the number of measurement sessions for each condition or with a single session (15 s of data). Figure 20 shows the obtained results. The data on this figure shows that the accuracy decreases when HoldPass is tested on unseen experimental conditions. But we also observe that this can be corrected by including a single measurement session from the target condition (15 s) in the startup phase of the system. HoldPass reaches an average accuracy of 84.9 % in that more challenging scenario.

We also performed similar cross conditions testing with different ways of holding/rotating the phone : rotation of 30° and 60° around X axis (tilting the phone towards the user) and rotation of 30° around Y axis (tilting the phone towards the ground). Results plotted on Figure 21 lead to similar observations while highlighting that this holding pattern has a less important effect on HoldPass' performance. With one 15 s-long session from each condition, HoldPass achieves an average accuracy of 87.96 % in that case.

8 LIMITATIONS AND DISCUSSION

In this section, we discuss the limitations of the current implementation of HoldPass and ways in which it could be improved as part of future work.

Heart condition changes. In our evaluation (Section § 7.7), we show that HoldPass performs well over time. This is because, as shown by prior works [38, 60], heart activity can be stable over time. Nevertheless, there are factors that can bring changes to heart activity, including heart disease, heavy exercising, etc. Any significant change in the heart activity pattern could negatively impact HoldPass's performance. Currently, all heart activity based authentication systems suffer this limitation. To cope with this limitation, HoldPass could be trained on data from these activities and/or perform a continuous updates of the user model based on Hand-BCG signals collected during different successful login phases. The latter can be achieved using online learning algorithms such as Stochastic Gradient Descent [5] which updates the model weights after processing each input sample. As a complement, in case of highly different inputs (multiple authentication failures), the system could ask the user if they wish to include these new data that might correspond to new usage conditions and validate their decision based on another authentication mechanism.

Remote model training. In the current design of HoldPass, we assume that the user model training is done offline on a remote server as this process generally requires significant computing capabilities. This can be seen

as a limitation since a spoofer might intercept the enrollment data over the network when they are sent to the server. To cope with this limitation, the data can be secured with strong encryption before being sent for model training on the remote server. A long-term solution is to prevent the data from being sent over the network at all and perform the model training directly on the user mobile phone. By exploiting recent advances in mobile machine learning, specifically the trend of on-device training [42, 63], HoldPass will eventually be able to perform the training or fine-tuning on the user device directly. This would at the same time facilitate the continuous update of the model to cope with the limitation presented in the previous paragraph.

Effect of motion. In the current design of HoldPass, the user is asked not to move when performing the authentication, or otherwise the system cannot reliably measure their Hand-BCG. Fortunately, HoldPass can achieve good accuracy with as little as 0.5 s of data (§ 7.8), mitigating this limitation. However, the system will not work in other scenarios such as on a moving vehicle, for example.

Advanced Replay Attacks. While HoldPass is by design robust against replay attacks as it is difficult for an external user to reproduce the hand vibrations of the legitimate user, an advanced malicious user can manage to hack the operating system and obtain the legitimate user's IMU data or trained model. The spoofed data can later on be injected directly as input to the login system which will therefore authorize access to the mobile phone. To prevent this type of attacks, HoldPass could perform liveliness detection by studying and exploiting the correlation of Hand-BCG with other sensors readings, e.g. heartbeat sound, which can be heard even in noisy environments as it mostly lies in an inaudible frequency band [25].

9 RELATED WORK

Authentication on Mobile Phones. Traditional user authentication on mobile phones are mostly based on passwords [46] or unlock patterns [44]. They provide simple, low-cost and easy-to-implement authentication, which however comes with the great limitation that they are sensitive to multiple types of replay attacks including smudge attacks and man-over-the-shoulder attacks [55]. To overcome this limitation, previous works exploit fingerprint [52], face [1], voice [16] and iris [33] to enable biometrics-based authentication on mobile phones. If largely adopted by the market, these approaches have been shown to be vulnerable to new types of replay attacks based on recorded or forged biometric inputs [14, 17, 21, 34, 56] which are moreover becoming increasingly sophisticated at lower cost.

In order to provide a simple second factor authentication, other works have also employed behavioral biometrics to authenticate the user while they type their password or draw their unlock pattern [9, 67], therefore enhancing the security of these traditional authentication modalities. Behavioral biometrics however have the well-known limitation that they are very prone to change over time [65].

New Sensing Modalities for Biometric Authentication. Because of the limitations of traditional biometrics based approaches, recent research works have investigated the use of other, and often new, biometric features for user authentication. SmileAuth [30] and BiLock [70] propose to authenticate the user based on information extracted from their teeth alignment or occlusion sound. VeinDeep [69] proposes to authenticate the user based on the vein pattern of their hand while BreathPrint [8] exploits audio features derived from an individual's breathing gestures for this task. If providing new interesting authentication modalities, these approaches still exhibit the important limitations that they either rely on specialized sensors that are not found on common mobile phones [69] or require a smiling or dental occlusion action, or explicit breathing with the phone near to the nose, that might not be convenient for use in public [8, 30, 70].

Building on top of an advanced litterature on heart activity unicity of a subject, established based on studies using specialized devices [4, 12, 18, 25, 27, 37, 43], other research works propose to authenticate the user either based on their Seismocardiogram (SCG) acquired with the phone placed on their chest or based on their Photoplestymogram (PPG) acquired with the phone placed on the phone camera. SCG-based authentication [60]

however has the limitation that it does not provide a satisfactory user experience as it requires the user to place the phone on their chest and PPG-based one does not provide enough distinctive information about the user and lead to an unsufficient accuracy in realistic cross-session authentication scenarios [41].

Ballistocardiography based Authentication. Since its beginning during the 19^{th} century [20, 51], Ballistocardiography has mostly been applied for diagnosis purposes, mostly focusing on Heart Rate and Heart Rate Variability monitoring [7, 11, 19, 29, 36, 58]. As this signal is related to heart activity which is a unique identifier of a human being, few recent research works have investigated its use for authentication purposes. [22] and [68] studied the possibility to authenticate a user based on 10 s-long whole body BCG signals collected with custom load cells sensors attached to a chair, while [24] proposes to use a Smart Eyewear to authenticate the user based on 3 s-long Head-BCG signals. If validating the possibility to use Ballistocardiography to authenticate users, these approaches are still not applicable to real-life mobile phones authentication as they require specialized hardware.

Vibration based Authentication. Different from BCG which is related to heart activity, other works propose to use vibration signals resulting from hand taping actions or vibration motor to authenticate the user. Taprint [10] exploits the vibration generated when a user taps on the fix knots of their hand to perform authentication on smartwatches. VibWrite [40] exploits the modification that a user finger incurs on a specific vibration signal to build a system that can be used on various planes. Following the same idea, VibID [66] and TouchPass [64] authenticate the user based on physical characters extracted from touching fingers when the phone[64] or smartwatch [66] emits a vibration. Tapprints [45] shows that the location of screen taps on smartphones and tablets can be identified from IMU readings. However, except TouchPass [64] which presents an interesting alternative, these works either require custom hardware (standalone vibration motor and piezoelectric sensor for VibWrite [40]) or are designed specifically for weareables [10, 45, 66], and are then not suitable for user authentication on COTS mobile phones.

Compared to all these works, HoldPass does not require any specialized hardware and runs in real time on commodity mobile phones with as short as 0.5 s of data to perform heart activity based authentication in a convenient to use manner.

10 CONCLUSIONS

In this paper, we presented HoldPass, the first system that can authenticate a user while they simply hold their phone, based on their unique hand vibration sensed with standard sensors found on commodity mobile phones. HoldPass addresses the unique challenges raised by properties of these Hand-BCG signals by introducing a novel alignment-free authentication scheme that builds on asynchronous signal slicing and a data-driven algorithm for identifying a reduced set of features for characterizing a user. We implemented HoldPass and evaluated its performance both with a large scale study involving more than 100 volunteers and targeted studies with smaller set of volunteers over a period of several months. The result showed that HoldPass authenticates the user with an accuracy and user experience similar to or better than state-of-the-art systems with stronger requirements on hardware and/or user participation.

ACKNOWLEDGMENT

This work was supported in part by the Agence Nationale de la Recherche under the ANR JCJC CiTADEL grant. Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (see https://www.grid5000.fr).

REFERENCES

[1] Paolo Abeni, Madalina Baltatu, and Rosalia D'Alessandro. 2006. *A Face Recognition System for Mobile Phones.* Vieweg, Wiesbaden, 211–217. https://doi.org/10.1007/978-3-8348-9195-2_23

[2] H. Aly and M. Youssef. 2016. Zephyr: Ubiquitous accurate multi-sensor fusion-based respiratory rate estimation using smartphones. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*. Institute of Electrical and Electronics Engineers, New York, NY, USA, 1–9. https://doi.org/10.1109/INFOCOM.2016.7524401

[3] Apache Commons Maths [n.d.]. Commons Math: The Apache Commons Mathematics Library. http://commons.apache.org/proper/commons-math/.

[4] Juan Arteaga-Falconi, Hussein Al Osman, and Abdulmotaleb El Saddik. 2015. ECG Authentication for Mobile Devices. *IEEE Transactions on Instrumentation and Measurement* 65 (12 2015), 1–10. https://doi.org/10.1109/TIM.2015.2503863

[5] Léon Bottou et al. 1991. Stochastic gradient learning in neural networks. *Proceedings of Neuro-Nımes* 91, 8 (1991), 12.

[6] L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen. 1984. *Classification and Regression Trees*. Taylor & Francis.

[7] C Brüser, Stefan Winter, et al. 2013. Robust inter-beat interval estimation in cardiac vibration signals. *Physiological measurement* (2013). https://doi.org/10.1088/0967-3334/34/2/123

[8] Jagmohan Chauhan, Yining Hu, et al. 2017. BreathPrint: Breathing Acoustics-Based User Authentication. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services* (Niagara Falls, New York, USA) *(MobiSys '17)*. ACM, New York, NY, USA, 278–291. https://doi.org/10.1145/3081333.3081355

[9] Huijie Chen, Fan Li, et al. 2020. Listen to Your Fingers: User Authentication Based on Geometry Biometrics of Touch Gesture. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 75 (Sept. 2020), 23 pages. https://doi.org/10.1145/3411809

[10] Wenqiang Chen, Lin Chen, Yandao Huang, Xinyu Zhang, Lu Wang, Rukhsana Ruby, and Kaishun Wu. 2019. Taprint: Secure Text Input for Commodity Smart Wristbands. In *Proceedings of the 25th Annual International Conference on Mobile Computing and Networking* (Los Cabos, Mexico) *(MobiCom '19)*. Association for Computing Machinery, New York, NY, USA, Article 17, 16 pages. https://doi.org/10.1145/3300061.3300124

[11] S.-T Choe and W.-D Cho. 2017. Simplified real-time heartbeat detection in ballistocardiography using a dispersion-maximum method. *Biomedical Research (India)* (2017).

[12] Tilendra Choudhary and M. Sabarimalai Manikandan. 2015. A novel unified framework for noise-robust ECG-based biometric authentication. In *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. 186–191. https://doi.org/10.1109/SPIN.2015.7095379

[13] Maximilian Christ, Nils Braun, Julius Neuffer, and Andreas W. Kempa-Liehr. 2018. Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing* 307 (2018), 72–77. https://doi.org/10.1016/j.neucom.2018.03.067

[14] Phillip L. De Leon, Michael Pucher, et al. 2012. Evaluation of Speaker Verification Security and Detection of HMM-Based Synthetic Speech. *IEEE Transactions on Audio, Speech, and Language Processing* 20, 8 (2012), 2280–2290. https://doi.org/10.1109/TASL.2012.2201472

[15] Helsinki Declaration. 2013. Ethical principles for medical research involving human subjects.

[16] Najim Dehak, Patrick J. Kenny, et al. 2011. Front-End Factor Analysis for Speaker Verification. *IEEE Transactions on Audio, Speech, and Language Processing* 19, 4 (2011), 788–798. https://doi.org/10.1109/TASL.2010.2064307

[17] Nesli Erdogmus and Sébastien Marcel. 2014. Spoofing Face Recognition With 3D Masks. *IEEE Transactions on Information Forensics and Security* 9, 7 (2014), 1084–1097. https://doi.org/10.1109/TIFS.2014.2322255

[18] Mohammed Farrag, M. Abo-Zahhad, et al. 2016. Heart-ID: Human Identity Recognition Using Heart sounds Based on Modifying Mel-Frequency Cepstral Features. *IET Biometrics* 5 (04 2016). https://doi.org/10.1049/iet-bmt.2015.0033

[19] Rafael Gonzalez-Landaeta, Oscar Casas, et al. 2008. Heart rate detection from an electronic weighing scale. *Physiological measurement* (2008). https://doi.org/10.1088/0967-3334/29/8/009

[20] JW Gordon. 1877. Certain Molar Movements of the Human Body produced by the Circulation of the Blood. *Journal of anatomy and physiology* (1877).

[21] Cisco Talos Group. 2020. *Fingerprint cloning: Myth or reality?* Retrieved Jul 19, 2021 from https://blog.talosintelligence.com/2020/04/fingerprint-research.html

[22] Hong Guo, Xinrong Cao, et al. 2013. Ballistocardiogram-based person identification using correlation analysis. In *World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China*, Mian Long (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 570–573.

[23] John E. (John Edward) Hall. 2016. *Guyton and Hall textbook of medical physiology* (13th ed.. ed.). Elsevier, Philadelphia, PA.

[24] Joshua Hebert, Brittany Lewis, Hang Cai, Krishna K. Venkatasubramanian, Matthew Provost, and Kelly Charlebois. 2018. Ballistocardiogram-based Authentication using Convolutional Neural Networks. arXiv:1807.03216 [eess.SP]

[25] Chenyu Huang, Huangxun Chen, et al. 2018. BreathLive: Liveness Detection for Heart Sound Authentication with Deep Breathing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 12 (March 2018), 25 pages. https://doi.org/10.1145/3191744

[26] O. T. Inan, P. Migeotte, et al. 2015. Ballistocardiography and Seismocardiography: A Review of Recent Advances. *IEEE J-BHI* (2015). https://doi.org/10.1109/JBHI.2014.2361732

[27] Mohit Ingale, Renato Cordeiro, et al. 2020. ECG Biometric Authentication: A Comparative Analysis. *IEEE Access* 8 (2020), 117853–117866. https://doi.org/10.1109/ACCESS.2020.3004464

[28] Steven A. Israel, John M. Irvine, et al. 2005. ECG to Identify Individuals. *Pattern Recogn.* 38, 1 (Jan. 2005), 133–142. https://doi.org/10.1016/j.patcog.2004.05.014
[29] Z. Jia, M. Alaziz, et al. 2016. HB-Phone: A Bed-Mounted Geophone-Based Heartbeat Monitoring System. In *ACM/IEEE IPSN*.
[30] Hongbo Jiang, Hangcheng Cao, et al. 2020. SmileAuth: Using Dental Edge Biometrics for User Authentication on Smartphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 84 (Sept. 2020), 24 pages. https://doi.org/10.1145/3411806
[31] Kevin Jiokeng, Gentian Jakllari, and André-Luc Beylot. 2021. *Hand-BCG & SCG signals dataset.* https://doi.org/10.5281/zenodo.5187910
[32] Chang-Sei Kim, Stephanie Ober, et al. 2016. Ballistocardiogram: Mechanism and Potential for Unobtrusive Cardiovascular Health Monitoring. *Scientific Reports* 6 (08 2016), 31297. https://doi.org/10.1038/srep31297
[33] Ajay Kumar and Arun Passi. 2010. Comparison and Combination of Iris Matchers for Reliable Personal Authentication. *Pattern Recogn.* 43, 3 (March 2010), 1016–1026. https://doi.org/10.1016/j.patcog.2009.08.016
[34] Sandeep Kumar, Sukhwinder Singh, and Jagdish Kumar. 2017. A comparative study on face spoofing attacks. In *2017 International Conference on Computing, Communication and Automation (ICCCA)*. 1104–1108. https://doi.org/10.1109/CCAA.2017.8229961
[35] M. Kyoso and A. Uchiyama. 2001. Development of an ECG identification system. In *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 4. 3721–3723 vol.4. https://doi.org/10.1109/IEMBS.2001.1019645
[36] F. Landreani, M. Morri, et al. 2017. Ultra-short-term heart rate variability analysis on accelerometric signals from mobile phone. In *EHB*.
[37] Ming Li and Xin Li. 2014. Verification based ECG biometrics with cardiac irregular conditions using heartbeat level and segment level information fusion. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 3769–3773. https://doi.org/10.1109/ICASSP.2014.6854306
[38] Feng Lin, Chen Song, et al. 2017. Cardiac Scan: A Non-Contact and Continuous Heart-Based User Authentication System. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking* (Snowbird, Utah, USA) *(MobiCom '17)*. ACM, New York, NY, USA, 315–328. https://doi.org/10.1145/3117811.3117839
[39] Wilburta Q Lindh, Marilyn Pooler, Carol D Tamparo, Barbara M Dahl, and Julie Morris. 2013. *Delmar's comprehensive medical assisting: administrative and clinical competencies*. Cengage Learning.
[40] 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 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA) *(CCS '17)*. Association for Computing Machinery, New York, NY, USA, 73–87. https://doi.org/10.1145/3133956.3133964
[41] Giulio Lovisotto, Henry Turner, et al. 2020. Seeing Red: PPG Biometrics Using Smartphone Cameras. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 3565–3574. https://doi.org/10.1109/CVPRW50498.2020.00417
[42] Jacopo Mangiavacchi and Santiago Castro. 2020. SwiftCoreMLTools: A Swift Library for creating CoreML models in Swift. https://github.com/JacopoMangiavacchi/SwiftCoreMLTools.
[43] Takhellambam Meitei, Ajit Sinam, and Swanirbhar Majumder. 2018. *PCG BASED BIOMETRIC*. 1–25. https://doi.org/10.4018/978-1-5225-5152-2.ch001
[44] Weizhi Meng, Wenjuan Li, et al. 2016. *On Multiple Password Interference of Touch Screen Patterns and Text Passwords*. ACM, New York, NY, USA, 4818–4822. https://doi.org/10.1145/2858036.2858547
[45] Emiliano Miluzzo, Alexander Varshavsky, Suhrid Balakrishnan, and Romit Roy Choudhury. 2012. Tapprints: Your Finger Taps Have Fingerprints. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services* (Low Wood Bay, Lake District, UK) *(MobiSys '12)*. Association for Computing Machinery, New York, NY, USA, 323–336. https://doi.org/10.1145/2307636.2307666
[46] Robert Morris and Ken Thompson. 1979. Password Security: A Case History. *Commun. ACM* 22, 11 (Nov. 1979), 594–597. https://doi.org/10.1145/359168.359172
[47] PaymentsJournal. 2020. *By 2024, How Many Smartphone Owners Will Use Biometrics?* Retrieved Jul 19, 2021 from https://www.paymentsjournal.com/by-2024-how-many-smartphone-owners-will-use-biometrics/
[48] T.W. Shen, W.J. Tompkins, and Y.H. Hu. 2002. One-lead ECG for identity verification. In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, Vol. 1. 62–63 vol.1. https://doi.org/10.1109/IEMBS.2002.1134388
[49] GIMA S.p.A. [n.d.]. *PM10 PALM ECG*. Contec Medical Systems Co., Ltd. Retrieved Nov 17, 2021 from https://www.gimaitaly.com/prodotti.asp?sku=33246&dept_selected=580&dept_id=5801
[50] Mark Spoonauer. 2017. *iPhone X Face ID Slower Than Touch ID (But There's a Fix)*. Retrieved June 14, 2021 from https://www.tomsguide.com/us/iphone-x-face-id-speed-up,news-26060.html
[51] Isaac Starr, A. J. Rawson, et al. 1939. Studies on the estimation of cardiac output in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; the ballistocardiogram. *American Journal of Physiology-Legacy Content* (1939). https://doi.org/10.1152/ajplegacy.1939.127.1.1
[52] Qi Su, Jie Tian, et al. 2005. A Fingerprint Authentication System Based on Mobile Phone. In *Audio- and Video-Based Biometric Person Authentication*. Springer Berlin Heidelberg, Berlin, Heidelberg, 151–159.
[53] Shan Suthaharan. 2016. *Decision Tree Learning*. Springer US, Boston, MA, 237–269. https://doi.org/10.1007/978-1-4899-7641-3_10

[54] Amirtahà Taebi, Brian Solar, et al. 2019. Recent Advances in Seismocardiography. *Vibration* (2019). https://doi.org/10.3390/vibration2010005

[55] Furkan Tari, A. Ant Ozok, and Stephen H. Holden. 2006. A Comparison of Perceived and Real Shoulder-Surfing Risks between Alphanumeric and Graphical Passwords. In *Proceedings of the Second Symposium on Usable Privacy and Security* (Pittsburgh, Pennsylvania, USA) *(SOUPS '06)*. ACM, New York, NY, USA, 56–66. https://doi.org/10.1145/1143120.1143128

[56] Dreamlab Technologies. 2020. *Attacking Biometric Systems with 3D Printing*. Retrieved Jul 19, 2021 from https://dreamlab.net/en/blog/post/attacking-biometric-systems-with-3d-printing-1/

[57] G. V. Trunk. 1979. A Problem of Dimensionality: A Simple Example. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-1, 3 (1979), 306–307. https://doi.org/10.1109/TPAMI.1979.4766926

[58] Antti Vehkaoja, Satu Rajala, et al. 2013. Correlation approach for the detection of the heartbeat intervals using force sensors placed under the bed posts. *Journal of medical engineering & technology* (2013). https://doi.org/10.3109/03091902.2013.807523

[59] Esra Vural, Steven Simske, and Stephanie Schuckers. 2013. Verification of individuals from accelerometer measures of cardiac chest movements. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*. 1–8.

[60] Lei Wang, Kang Huang, et al. 2018. Unlock with Your Heart: Heartbeat-Based Authentication on Commercial Mobile Phones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 140 (Sept. 2018), 22 pages. https://doi.org/10.1145/3264950

[61] Wikipedia. [n.d.]. *Pegasus Project (investigation)*. Retrieved August 14, 2021 from https://en.wikipedia.org/wiki/Pegasus_Project_(investigation)

[62] Ian H. Witten, Eibe Frank, et al. 2016. *Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques* (4th ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[63] Yawen Wu, Zhepeng Wang, et al. 2020. Enabling On-Device CNN Training by Self-Supervised Instance Filtering and Error Map Pruning. *CoRR* abs/2007.03213 (2020). arXiv:2007.03213 https://arxiv.org/abs/2007.03213

[64] Xiangyu Xu, Jiadi Yu, et al. 2020. TouchPass: Towards Behavior-Irrelevant on-Touch User Authentication on Smartphones Leveraging Vibrations. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking* (London, United Kingdom) *(MobiCom '20)*. ACM, New York, NY, USA, Article 24, 13 pages. https://doi.org/10.1145/3372224.3380901

[65] Roman Yampolskiy and Venu Govindaraju. 2008. Behavioural biometrics: A survey and classification. *International Journal of Biometrics* 1 (01 2008). https://doi.org/10.1504/IJBM.2008.018665

[66] Lin Yang, Wei Wang, and Qian Zhang. 2016. VibID: User Identification through Bio-Vibrometry. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. 1–12. https://doi.org/10.1109/IPSN.2016.7460725

[67] Xinchen Zhang, Yafeng Yin, et al. 2020. TouchID: User Authentication on Mobile Devices via Inertial-Touch Gesture Analysis. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 162 (Dec. 2020), 29 pages. https://doi.org/10.1145/3432192

[68] X. Zhang, Y. Zhang, L. Zhang, H. Wang, and J. Tang. 2018. Ballistocardiogram Based Person Identification and Authentication Using Recurrent Neural Networks. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. 1–5. https://doi.org/10.1109/CISP-BMEI.2018.8633102

[69] Henry Zhong, Salil S. Kanhere, and Chun Tung Chou. 2017. VeinDeep: Smartphone unlock using vein patterns. In *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 2–10. https://doi.org/10.1109/PERCOM.2017.7917845

[70] Yongpan Zou, Meng Zhao, et al. 2018. BiLock: User Authentication via Dental Occlusion Biometrics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 152 (Sept. 2018), 20 pages. https://doi.org/10.1145/3264962