



# User Authentication Using Body Vibration Characteristics

Y. Zhang<sup>(✉)</sup> and Y. Ren

University of Electronic Science and Technology of China, No. 2006, Xiyuan Avenue,  
High tech Zone (West Zone), Chengdu, People's Republic of China  
zhangyx.123@foxmail.com, renyanzhi05@uestc.edu.cn

**Abstract.** Traditional mobile phone authentication systems are based on knowledge or biological information. This paper shows an authentication system based on human vibration characteristics and implement it on smartphone. This kind of inspiration comes from signal transmitted based on solid conduction and is applied to the link between vibration motor and accelerometer. Therefore, we designed a system to extract user biometric features by active vibration. However, there is a great challenge for smartphone systems with low sampling rate and requiring real-time response. Therefore, we would realize the system through multiple signal processing stages rather than choose a high-performance neural network. Our system solves the problem of low sampling rate of mobile phone sensors through supersampling reconstruction method. Besides, we select appropriate statistical features and MFCC-based features through PCA algorithm, and finally complete the training through Gradient Boosting Tree. In order to avoid the threshold division problem of the multi-level classifier, we train each sample in two classifications at the time of registration, and store the parameters in the user profile. When the system performs user authentication, the user data is divided into five sections for testing, so as to increase the robustness of the system. Our approach could achieve short-time identity authentication, with an average accuracy rate of 85.3%.

**Keywords:** Vibration · Authentication · Mobile Sensing

## 1 Introduction

The convenience of mobile devices such as smartphones and smartwatches has greatly stimulated the development of the mobile industry in recent years. However, although users can enjoy the great convenience brought by mobile devices, the widespread use of mobile devices and mobile applications has caused major security problems [1]. For example, most applications need to obtain permissions such as user location and phone information. Therefore, identification and verification before using mobile devices has become the first barrier to protect data security.

Nowadays, the identity authentication process of smart phones mainly involves some traditional solutions, including PIN, fingerprint identification, face recognition, etc. In the U.S. consumer payment study, 66% of users set PINs and passwords as the first choice in smartphone authentication. Although PINs based authentication methods are easy to use and widely deployed, they also have the problem of password leakage. For example, people have been able to crack it through shoulder-surfing [2]. Besides, many studies show that attackers can infer your mobile phone’s PIN through wifi signals [3]. Although the decryption cost of attackers will increase with the increase of PIN’s complexity, it increases the burden of users’ memory.

In addition, attackers can forge users’ physiological and biometric information (such as fingerprints and faces) to deceive the system. If the user loses such information, authentication based on physiological biometrics will be permanently insecure. For example, researchers found that fingerprint based authentication security may suffer from smudge attacks [4] and the attacker can spoof face recognition by 3D masks using micro-texture analysis [5,6]. Therefore, we need to design a new biometric-based identity authentication method.

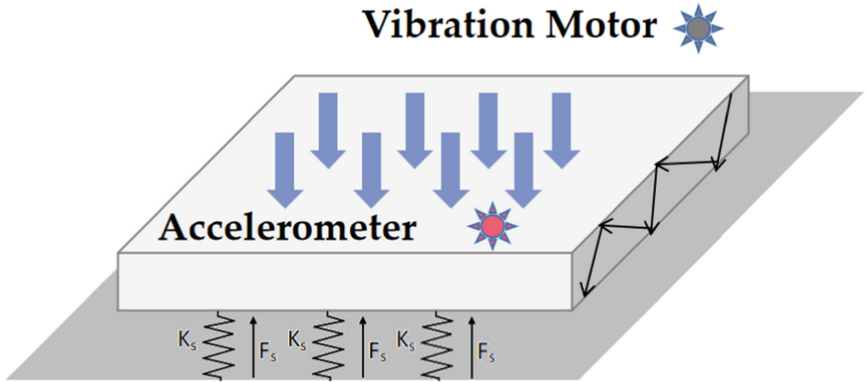
Moreover, high accuracy means expensive sensor costs. For most smartphone manufacturers, the hardware cost of smartphones is also an important fact to be considered. High performance sensors not only increase the cost of mobile phones, but also take up a lot of internal space of mobile phones. For example, the iris reader [7] of Samsung smartphones and the depth camera [8] of iPhone are expensive and vulnerable, so it is particularly critical to find an alternative authentication method.

Based on the existing research results, this paper focuses on the biological characteristics of the human body stimulated by vibration signals, and completes the training of the identity verification system by filtering and feature extraction of vibration signals. The main contributions of this work are summarized as follows: (1) Exploring the feasibility of smart phone authentication through the accelerometer and the vibration motor. (2) We analyze the signals with different vibration frequencies, study their influence on feature selection, and provide solutions to meet the needs of most existing mobile phone hardware devices. (3) For different environmental conditions, we propose an optimization algorithm to reduce the interference of noise to the system, so as to improve the stability of the system.

## 2 Background

We think that the vibration motor and accelerometer of the mobile phone work together as a system. The vibration wave generated by the vibration motor propagates through the surface of the mobile phone and is received by the accelerometer [9]. When propagating, when the vibration wave meets two different media boundaries, the vibration wave will form energy attenuation and multipath interference. Figure 1 shows reflection and diffraction of a vibration signal propagating

on a solid surface. Since the transmitted vibration signal reaches the accelerometer through reflection and diffraction, the accelerometer will have unique vibration characteristics (such as wave attenuation and multipath interference), so it can be used to identify intelligent devices [10].



**Fig. 1.** Vibration signal propagation.

Figure 1 shows the force condition of the mobile phone screen when the vibration motor is working. When the user touches the mobile phone with his hand, a downforce shock wave is generated on the mobile phone screen, which affects the propagation path of the vibration signal. Where  $k_s$  is the effective spring constant and  $k_d$  is the damping coefficient. If the vertical displacement of the surface is  $x$ , we have

$$F_t = K_d \left( \frac{d}{dt} \right) x + K_s x + M \left( \frac{d}{dt} \right)^2 x \quad (1)$$

This indicates that the finger touching force could be captured by analyzing the received vibration signals and utilized as a biometric-associated feature in our system.

In addition, some experiment demonstrate that the vibration energy absorbed into the human finger-hand-arm system is different under different vibration frequencies [11]. Therefore, we will explore the impact of vibration frequency on the authentication system in the following sections.

### 3 System Overview

In this section, we introduce a verification method based on human biological characteristics corresponding to vibration. As shown in Fig. 2, our system solves the problem of low sampling rate of mobile phone sensors through supersampling reconstruction method. Besides, we select appropriate statistical features and

MFCC-based features through PCA algorithm, and finally complete the training through Gradient Boosting Tree. In order to avoid the threshold division problem of the multi-level classifier, we train each sample in two classifications at the time of registration, and store the parameters in the user profile. When the system performs user authentication, the user data is divided into five sections for testing, so as to increase the robustness of the system.

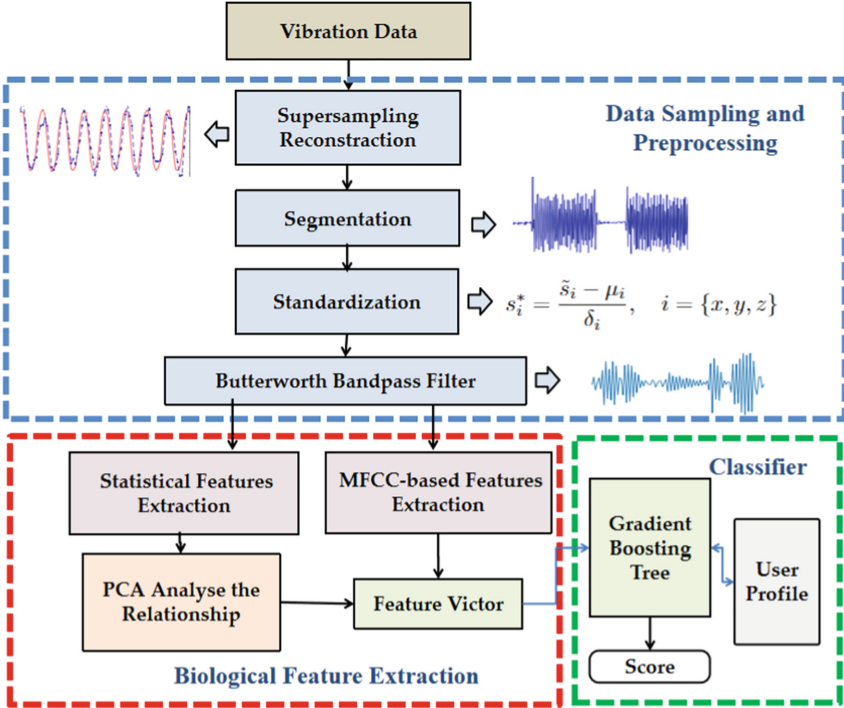


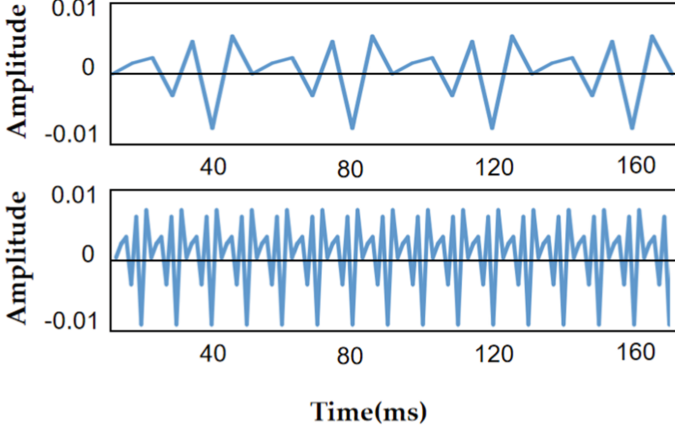
Fig. 2. System Overview.

### 3.1 Data Sampling and Preprocessing

**Supersampling Reconstruction Method.** According to the Nyquist Sampling Law, the low sampling rate results in the distortion of vibration waveform in the time-domain. Because we need to use the amplitude peak value as the signal characteristic in the subsequent work, this will lead to the increase of measurement error. The accelerometer in iPhone 7 supports the maximum sampling rate 100 Hz, which is much less than the frequency of vibro-motor at 167 Hz [12]. So we have to adopt the sampling rate up 400 Hz.

Due to realize this supersampling reconstruction method(SSR), the signal needs to have sufficient stability. For example, as shown in Fig. 3, put two signals with different resolutions on the same time axis, and find the sum of the

minimized variances of the corresponding points. After calculation, the sum of the minimized variables of the corresponding points is extremely small. This means that we can obtain 400 Hz signal through the supersampling construction method



**Fig. 3.** Comparison of different frequency signals.

The specific method is that we can sample the same value instead of recording all signals for the current signal with too low sampling rate [13]. For example, if different sampling points are sampled in each cycle and timestamps are recorded, there will be a large number of labeled sampling points after several cycles. Next, we will combine them into a complete cycle and sort them.

But the more complicated problem is the determination of the sampling interval. It can be seen from the Fourier transform formula that,

$$X(e^{j\omega}) = \sum_{n=1}^N x[t_n] e^{-j\omega t_n} \quad (2)$$

When  $t_n$  is replaced by an arbitrary random number, the discrete fourier transform will introduce random noise in the frequency domain. Here we assume that the time  $t_n$  follows a uniform distribution, and the expectation of the spectrum can be obtained as follows:

$$\begin{aligned} E[X(e^{j\omega})] &= \frac{1}{T_{\max}} \sum_{n=1}^N \int_0^{T_{\max}} x[t_n] e^{-j\omega t_n} dt_n \\ &= \frac{N}{T_{\max}} X(j\omega) \end{aligned} \quad (3)$$

In our system, we set the cycle composition of 0.5 s active vibration and  $t_{\text{gap}}$  to 5 ms to apply SRR for four cycle reconstruction.

**Standardization and Filtering.** When the smartphone’s motor vibrates, the accelerometer generates a specific feedback signal. In the data collection part, we need to preprocess the  $acc_x$ ,  $acc_y$  and  $acc_z$  data obtained from the accelerometer, aiming at removing high-frequency noise, and normalizing and aligning the signals to ensure the system robustness under different postures.

Coordinate system modification. In general, when a user authenticates a smartphone, there is no guarantee that the user can maintain the absolute level of the smartphone. Therefore, the built-in accelerometer of the smartphone makes a huge difference in each verification process. In order to ensure that our equipment can operate stably in various environments, we need to correct some data of the coordinate system. We subtract the gravitational acceleration from the projection of the accelerometer on the three coordinate axes, and pass the low-pass filter [14].

$$\tilde{s}_i = (1 - \beta)(s_i - g_i), \quad i = \{x, y, z\} \quad (4)$$

$$\beta = \frac{dT}{t + dT} \quad (5)$$

where  $g_i$  and  $s_i$  are the projection of the gravitational acceleration and raw acceleration captured by the accelerometer along the  $i$ -th axis, respectively;  $\tilde{s}_i$  is the associated acceleration after such an alignment;  $\beta$  is a filter factor determined by filter’s time constant  $t$  and event delivery rate  $dT$ . In this work, we empirically choose  $\beta$  to be 0.2.

The accelerations and angular velocities collected by accelerometers and gyroscopes differ greatly among the three directions, even more among different device models. To ensure the numerical comparability and analysis stability, our system applies the  $Z_{score}$  standardization method [15] to the readings from each axis as follows:

$$s_i^* = \frac{\tilde{s}_i - \mu_i}{\delta_i}, \quad i = \{x, y, z\} \quad (6)$$

where  $\tilde{s}$  is a single reading along the  $i$ -th axis after filtering,  $\mu_i$  and  $\delta_i$  are the mean and standard deviation of all  $\tilde{s}$  along the same axis respectively. After the standardization,  $s_i^*$  is centered at 0 and scaled to have the standard deviation of 1. See Fig. 4 for the normalized signal.

For noise interference brought by the environment, such as music and thermal noise, and interference, such as arm movement and shaking, we choose a low-pass filter to reduce these effects. Through analysis, it can be found that the frequency of the built-in vibration motor of existing smartphones is generally between 150–250 Hz, while the motion frequency of humans is 10 Hz [16]. Therefore we develop a Butterworth bandpass using the cutting-off frequencies 10 Hz 250 Hz to filter the vibration noises and interferences outside this range. The filtered signal image is shown in Fig. 4.

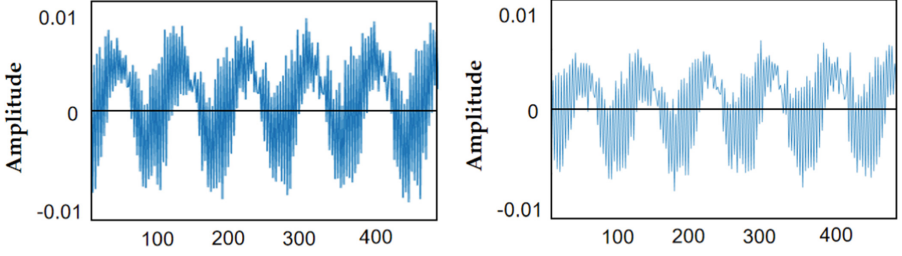


Fig. 4. Comparison before and after standardization and filtering.

### 3.2 Feature Extraction

**MFCC-based Feature Extraction.** The Mel-frequency cepstral coefficient (MFCC) is widely used to represent the short-term power spectrum of acoustic or vibration signals [17] and can represent the dynamic features of the signals with both linear and nonlinear properties. While the MFCCs are able to distinguish people’s sound differences in speech and voice recognition, we find that they can also characterize the vibration signals transmitting via the medium of a solid surface on which the user’s finger touches [18]. The MFCC feature extraction process is shown in the Fig. 5, mainly including pre emphasis, framing, windowing, fast Fourier transform (FFT), Mel filter bank, discrete cosine transform (DCT). Among them, FFT and Mel filter bank are the most important.

$$mfcc(i, n) = \sum_{m=1}^M \log[H(i, m)] \cdot \cos \left[ \frac{\pi \cdot n \cdot (2m - 1)}{2M} \right] \quad (7)$$

$$K(i) = 1 + \left( \frac{L}{2} \right) \cdot \sin \left( \frac{\pi \cdot i}{L} \right) \quad i = 1, 2, 3, \dots, 13 \quad (8)$$

where  $M$  represents the number of Mel filters,  $i$  represents the data of the  $i$ -th frame, and  $n$  represents the  $n$ -th column of the  $i$ -th frame (the value range of  $n$  is 1–26). In our system, we calculate the MFCCs of each segment of signal. We set 26 Mayer filters and calculate in each 50 ms the Hamming window to

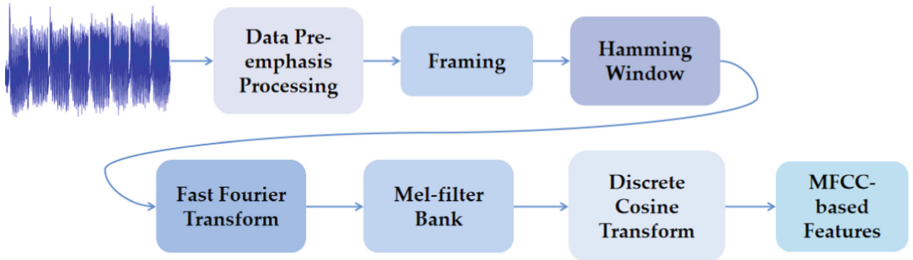


Fig. 5. MFCC Flowchart.

obtain 26 zero order MFCC eigenvalues. Because most of the signal data is generally concentrated in the low-frequency region after conversion, only the first 13 data are taken as MFCC based features for each frame. After MFCC feature extraction of 5S vibration signal, we can get a  $11 \times 13$  MFCC-based feature matrix.

**Statistical Features Extraction.** Although MFCC-based feature has a good classification effect in high-frequency signals, it is not good in low-frequency signals. This makes it particularly difficult to classify different users simply through MFCC-based features. Therefore, we need to obtain more information of the low-frequency part from the time-domain and frequency-domain of the signal, considering of statistical features along with pairs of peak indicators and heights in the frequency domain and in the correlation of time domain.

**Table 1.** The total of 32 features for each response signal.

Abbreviation	Explanation
Var	variance
MAV	mean absolute value
RMS	root mean square
Std	standard deviation
IOR	interquartile range
Energy	Integral of square of signal amplitude
Entropy	Shannon entropy of continuous time series
Correlation	pairs of indices and heights of the highest five peaks
FFT	pairs of indices and heights of the highest five peaks
DCT	pairs of indices and heights of the highest five peaks
DWT	pairs of indices and heights of the highest five peaks
PSD	pairs of indices and heights of the highest five peaks

In the time domain, the statistical features are variance (Var); mean absolute value (MAV); root mean square (RMS); standard deviation (Std); interquartile range (IQR); energy; entropy; pairs of indices and heights of the highest five peaks in the correlation. Also, in frequency domain, we extract pairs of indices and heights of the highest five peaks after using fast fourier transform (FFT), discrete cosine transform (DCT), discrete wavelet transform (DWT), and power spectral density (PSD). A total of our features is shown in Table 1. We have a total of 32 features for each response signal.

We tested 5 experimenters and analyzed the test data by principal component analysis (PCA). The Fig. 6 indicated that at the sampling rate 400 Hz, statistical features are associated but MFCC-base features are loosely associated. Therefore, it is necessary to combine them as input features of classifier.

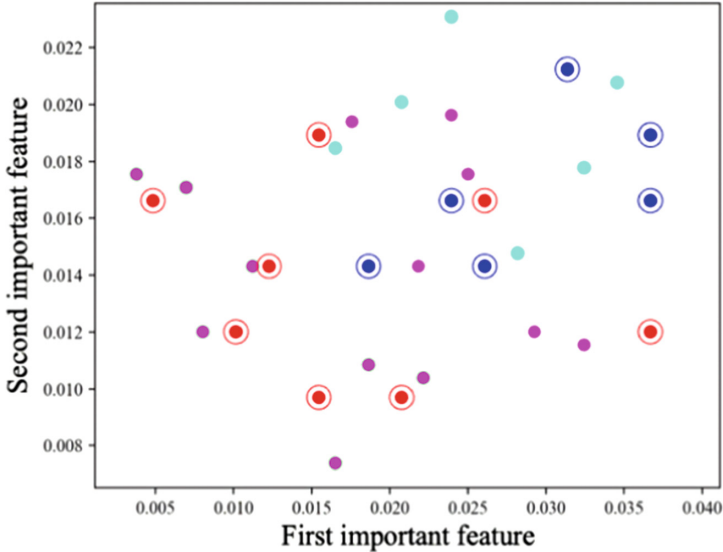


Fig. 6. User classification based on MFCC features and statistical features.

### 3.3 User Classification

We build a binary classifier for each user by using the Gradient Boosting Tree (GBT). We choose GBT mainly because (1) GBT is famous for its robustness to various types of features with different scales, which is the exact case in our project (e.g., the energy of the vibration signal is around 5, and the coefficients are the numbers fluctuated around 0 with value less than 1). Therefore, GBT would eliminate the efforts to normalize or whiten the feature data before classification. (2) GBT classifier is robust to the collinearity of feature data. Because our features are heterogeneous across different domains, it may result in unexpected correlation or unbalance ranges that possess the collinearity [19]. This means that we do not need to analyze the correlation of features, thus reducing the complexity of the algorithm.

Given  $N$  training samples  $(x_i, y_i)$ , where  $x_i$  and  $y_i$  represent the feature vector (including statistical features and MFCC-based features) and corresponding user label (i.e.,  $y_i = 1$  or 0 represents whether  $x_i$  is from corresponding user), GBT seeks a function to iteratively select weak learners  $h_j$  and their weight  $\omega_j$  to minimise the loss function [20].

$$\phi(x_i) = \sum_{m=1}^M \omega_m h_m(x_i) \quad (9)$$

We adopt the GBT implementation from the library of SQBlib, such as enough shrinkage (i.e., 0.1) and number of iterations (i.e.,  $M = 2000$ ). The above parameters adopted in GBT are optimized in terms of the speed and accuracy based on our empirical study.

When we registering a new user registers, the system will extract features from the segmented samples, and then input these features into GBT for training. During the training, the target user is marked as 1, and other users are marked as 0. After the training is completed, the parameters can be stored locally. In the user verification link, we divided the user’s data within 2.5s into five parts, and each part was tested separately. Each binary gradient classifier will output a score for the testing feature set. Finally, we calculate the score for  $i$ -th classifier through these five segmentations, and outputs the user ID corresponding to the maximum value.

$$\text{Score}_i = \frac{1}{5} (S_1 + S_2 + S_3 + S_4 + S_5) \quad (10)$$

$$\text{Output} = \max(\text{Score}_1, \text{Score}_2, \text{Score}_3, \text{Score}_4, \dots, \text{Score}_i) \quad (11)$$

## 4 Experimental Setup

### 4.1 Environment

To apply our vibration-based user authentication method, we use the iPhone 7, which can represent the basic performance of most mobile phones at present. The vibration frequency of vibro-motor was 167 Hz, and the accelerometer’s sampling rate 100 Hz. The difference in waiting time was 50 ms, and we stopped sampling after the motor restarts five times.

### 4.2 System Performance

Here, we utilize the false rejection rate (FRR) and false acceptance rate (FAR) as metrics to evaluate the authentication accuracy of our system. FAR is the fraction of other users’ data that are misclassified as the legitimate user’s. FRR is the fraction of the legitimate user’s data that are misclassified as other users’ data. For security protection, a large FAR is more harmful than a large FRR. However, a large FRR would degrade the usage convenience.

To verify the effectiveness of our proposed model and techniques, we first collected 50 sets of data on a stationary desktop, and 50 sets of data during hand lifting from 3 experimenters, and the data time of each group was 2.5s, forming a total of 800 samples. We utilized 10 fold cross validation for training and testing, and obtained the results shown in the Table 2.

**Table 2.** The FRR, FAR and accuracy of system.

State	FRR	FAR	Accuracy
Static	13.3%	8.5%	90.1%
Movement	17.5%	11.9%	85.3%

To explore the relationship between sample length and accuracy, we tested the changes of FRR and FAR under different sample lengths. The results are shown in the Fig. 7.

We found that with the increase of the sample length, the accuracy of the verification continued to rise. However, since the system needs to provide a better user experience, we believe that when the sample length is greater than 2.5 s, the small improvement in accuracy obtained by increasing the length is not cost-effective. In addition, with the increasing sampling rate of mobile phone sensors, the final accuracy of the system is also improving. This means that the technology has a higher upper limit in the future.

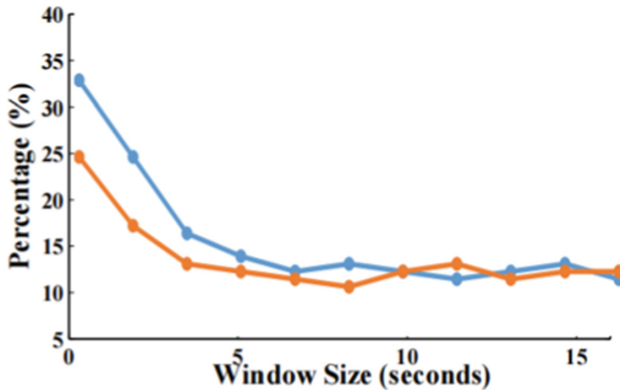


Fig. 7. FRR and FAR under different sample lengths.

## 5 Conclusion

In this paper, we proposed a vibration-based user authentication method for smartphone, which does not require user's personal information or privacy. We evaluated our method on a commercial smartphone, the iPhone 7, and default vibration types officially provided, which means no additional devices are required to authenticate users. In addition, our method produced a low EER of 0.147 for short-term signals. We expect our method to be suitable for a wide variety of smartphone on the market today.

## References

1. Mayrhofer, R., Gellersen, H.: Shake well before use: intuitive and secure pairing of mobile devices. *Mob. Comput. IEEE Trans.* **8**(9), 729–806 (2009)
2. Chong, M.K., Mayrhofer, R., Gellersen, H.: A survey of user interaction for spontaneous device association. *ACM Comput. Surv. (CSUR)* **47**(1), 8:1–8:40 (2014)

3. Li, M., et al.: When CSI meets public WiFi: inferring your mobile phone password via WiFi signals. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ACM, 2016, pp. 1068–1079 (2016)
4. Von Zezschwitz, E., Koslow, A., De Luca, A., Hussmann, H.: Making graphic-based authentication secure against smudge attacks. In: Proceedings of the 2013 International Conference on Intelligent User Interfaces, ACM, 2013, pp. 277–286 (2013)
5. Määttä, J., Hadid, A., Pietikäinen, M.: Face spoofing detection from single images using micro-texture analysis. In: Biometrics (IJCB), 2011 International Joint Conference on, IEEE, 2011, pp. 1–7 (2011)
6. Erdogmus, N., Marcel, S.: Spoofing face recognition with 3D masks. *IEEE Trans. Inf. Forensics Secur.* **9**(7), 1084–1097 (2014)
7. Galleso, M.: Samsung, Galaxy Note 7: An Easy Guide to the Best Features, First Rank Publishing, 20 October 2016
8. Bud, A.: Facing the future: the impact of Apple FaceID. *Biom. Technol. Today* **12018**(1), 5–7 (2018)
9. Cheng, J.Q., et al.: WatchID: wearable device authentication via reprogrammable vibration. In: Hara, T., Yamaguchi, H. (eds.) *Mobile and Ubiquitous Systems: Computing, Networking and Services. MobiQuitous 2021. LNICS, Social Informatics and Telecommunications Engineering*, vol. 419. Springer, Cham (2022). [https://doi.org/10.1007/978-3-030-94822-1\\_53](https://doi.org/10.1007/978-3-030-94822-1_53)
10. Tarzia, S.P., Dinda, P.A., Dick, R.P., Memik, G.: Indoor localization without infrastructure using the acoustic background spectrum. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (ACM MobiSys) (2011)
11. Chang, C.C., Lin, C.J.: 2011. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**(3), 1–27 (2011). <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
12. Yu, Y., et al.: Initial validation of mobile-structural health monitoring method using smartphones. *Int. J. Distrib. Sens. Netw.* **11**(2), 274391 (2015)
13. Huang, Y., et al.: Vi-liquid: unknown liquid identification with your smartphone vibration. In: Proceedings of the 27th Annual International Conference on Mobile Computing and Networking (2021)
14. AndroidDeveloper, Work with raw data, use the accelerometer. <https://developer.android.com/studio>
15. Kreyszig, E.: *Advanced engineering mathematics*, 10th edn. (2009)
16. Antonsson, E.K., Mann, R.W.: The frequency content of gait. *J. Biomech.* **18**(1), 39–47 (1985)
17. Murty, K.S.R., Yegnanarayana, B.: Combining evidence from residual phase and MFCC features for speaker recognition. *IEEE Signal Process. Lett.* **13**(1), 52–55 (2006)
18. Liu, J., et al.: Vibwrite: towards finger-input authentication on ubiquitous surfaces via physical vibration. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (2017)
19. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning*, pp. 115–163. Springer, New York, NY (2001). <https://doi.org/10.1007/978-0-387-84858-7>
20. Galar, M., Fernández, A., Barrenechea, E., Bustince, H., Herrera, F.: An overview of ensemble methods for binary classifiers in multi-class problems: experimental study on one-vs-one and one-vs-all schemes. *Pattern Recogn.* **44**(8), 1761–1776 (2011)