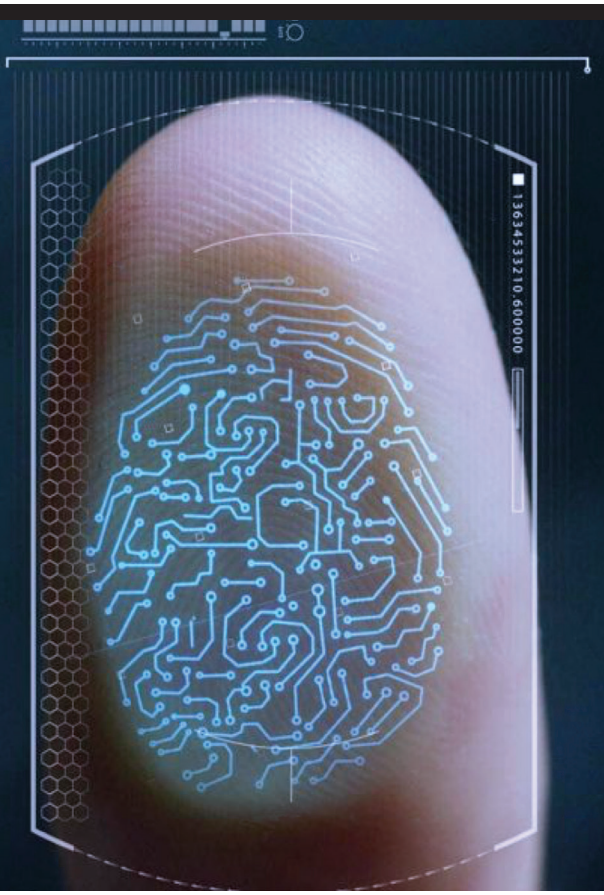
**Aditya Singh Rathore, Chenhan Xu and Wenyao Xu** *University at Buffalo, The State University of New York, Buffalo, NY*

**Editors: Nic Lane and Xia Zhou**



# SonicPrint: Discovering the Voice of Fingerprint for Adoptable Biometrics

Excerpted from "SonicPrint: a generally adoptable and secure fingerprint biometrics in smart devices," from *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services (MobiSys '20)*, with permission. https://dl.acm.org/doi/10.1145/3386901.3388939 © ACM 2020

Although fingerprint technology holds great promise for user authentication, commercial scanners face significant challenges in terms of security (e.g., fake finger) and adoptability (e.g., wearables). *SonicPrint* pushes the boundary of fingerprint biometrics beyond smartphones to any smart devices without the need for specialized hardware. To achieve this, it listens for fingerprint-induced sonic effect (*FiSe*) caused when a user swipes his/her fingertip on smart device surface. Compared to other biometrics including physiological patterns and passive sensing, SonicPrint is a low-cost, privacy-oriented and secure approach to identify users across smart devices of unique form-factors.

## INFORMATION AT YOUR FINGERTIPS

What if accessible surfaces encountered during day-to-day activities could sense the fingerprint information? Imagine a daily routine where Alice arrives in her home from the workplace. Facing the entrance, Alice swipes her fingertip on the metal door lock upon which she is granted permission to her house. Recognizing that it is Alice who performed the swipe action, the lights in the house are turned on and Alice's favorite music starts to play. Alice approaches the IoT sensors in her house and swipes anywhere on the sensor's plastic surface to once again verify her identity. This lets her adjust the surroundings to an ideal temperature. To Alice's surprise, Google Echo informs her about the unpaid electricity bills which she can pay either by swiping on her smartwatch or on the fiber surface of Google Echo. During every critical task, Alice has the freedom to swipe on any surface near the target device to verify her identity and ensure that only she can make decisions. To make this scenario a reality, where accessible surfaces can scan fingerprint information without the need for commercial scanners (see Figure 1), we build SonicPrint to leverage the intrinsic fingerprint ridge information in sonic waves for user identification.

## ACOUSTICS RATHER THAN VISUAL

We have taken a radically different approach to fingerprint sensing. Instead of trying to obtain the visual perception of fingerprint, we focus on its acoustic capabilities by leveraging the friction principles of sliding surfaces. Specifically, friction leads to distinct waves and oscillations within the interacting mediums resulting in the emission of sonic waves to the ambient environment [1]. Our key contribution is the observation that the sonic waves from a user swiping his fingertip on a surface can serve as biometric traits. Since every person has a unique fingerprint, two users swiping their fingertips on a common surface result in distinct fingerprint-induced

Photo, shutterstock.com

sonic effects. FiSe inherent uniqueness is dependent on the surface texture (i.e., fingerprint ridge patterns) and the finger's constitution while its audibility depends on the texture of the interacting surface. For instance, a user swiping on a coarse paper surface would have a higher sound pressure level than when the user swipes on a smooth silicon surface. Yet FiSe can be observed across common materials and measured using inbuilt microphones.

## MODELING THE FINGERPRINT'S VOICE

A human fingerprint can be visually perceived while FiSe lies in the audio domain. To retrieve the fingerprint-dependent characteristics from FiSe, we need to understand the different levels of information provided by a fingerprint (see Figure 2) and its semantic relationship to acoustic features.

- **Level I:** The macro-details of a fingerprint, such as patterns and ridge flows, can be seen through naked eyes. Technically, these features are characterized by local orientations (i.e., angle of the ridge with horizontal axis) and local frequencies (i.e., number of ridges per unit length) [2]. The pattern exhibits regions where ridge lines take distinctive shapes (whorl, arch, loop), which can be common among different users. Yet Level I features are highly intuitive since anyone can perceive the information. Similarly, in the audio domain, power-based temporal features that highlight changes over time and perceptual features, e.g., pitch, hold an intuitive meaning to a human listener and can provide high-level information of sonic waves.

- **Level II:** At the local scale, discontinuous ridges, commonly known as minutiae, can be observed in fingerprint patterns. For instance, a ridge can divide into two (bifurcations), abruptly end (termination) or form a unique shape (hook). Unlike Level I, these features possess high variance between different users and are actively used for authentication. Considering that minutiae textures affect the timbral characteristics of FiSe, cepstral features are essential to capture this influence and discriminate against the audio sources.


**FIGURE 1.** SonicPrint is a fingerprint sensing dimension that is adoptable across diverse smart devices and resilient to fake-finger spoofing.

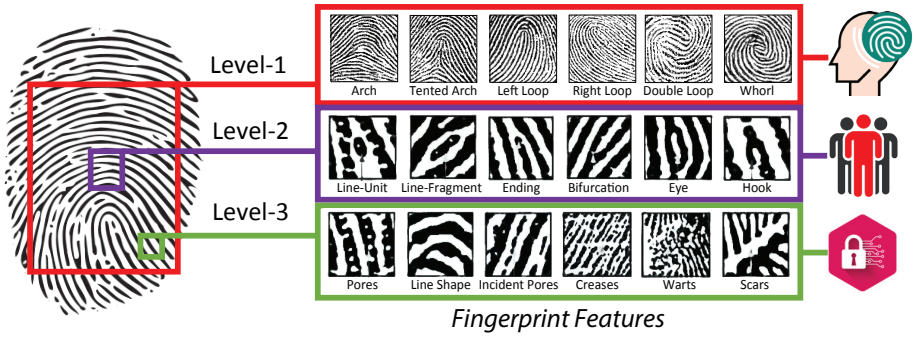



**FIGURE 2.** Fingerprint with three level characteristics focusing on perception, uniqueness and anti-spoofing respectively.

**TABLE 1.** Comparison of SonicPrint with widely used biometric techniques.

| Modality | Degree-of-freedom | Adoptability | Cost | Robustness | Security |
|---|---|---|---|---|---|
| **Face** | Medium | Medium | Low | Medium | Medium |
| **Iris** | Low | Low | Medium | High | Medium |
| **Voice** | Medium | Medium | Low | Medium | Low |
| **Fingerprint** | High | Medium | Low | Medium | Low |
| **SonicPrint** | High | High | Low | Medium | High |

- **Level III:** Although unique, Level II fingerprint features are prone to spoofing since they could still be visually perceived through the naked eye or even in low-resolution images. Thus, Level III features are proposed based on the dimensional ridge information, including width, pores and edge contour. Similarly, short-time Fourier transform and adaptive time-frequency decomposition can reveal various physical attributes of FiSe. These features have inferior meaning to human perception [3] and thus are difficult to spoof.

In our study, the Level I, II, III information is extracted from FiSe and utilized for user identification. With collaborative efforts and interdisciplinary research, further discoveries can be foreseen as researchers derive new taxonomies for modeling the acoustic fingerprint.


**FIGURE 3.** SonicPrint has a three-step deployment, with downloading the software solution, registering the biometric template and performing swipe actions, to unlock the smartphone.

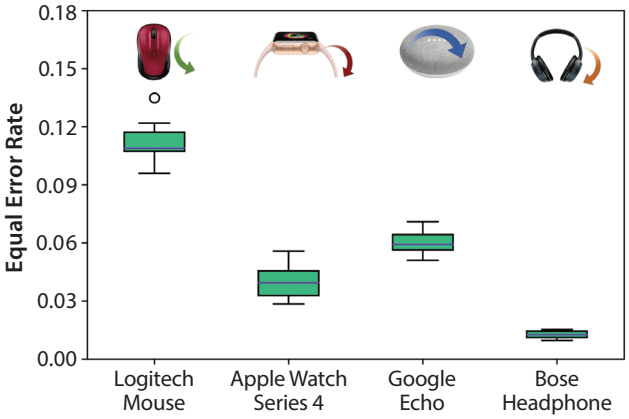



**FIGURE 4.** Evaluation of curved smart devices.

## IDEAL CHARACTERISTICS OF SONICPRINT

We believe that this novel method makes fingerprint sensing even more intuitive and transparent than recent ultrasound-based sensing. It embodies the following characteristics that are the foundation of a practical biometric:

**(i)** It does not require any specific hardware and utilizes low-cost off-the-shelf sensors in smart devices.

**(ii)** The biometric trait is available across devices and materials with diverse flexibility, geometry and composition.

**(iii)** It enables real-time data collection, noise elimination, feature engineering and training/inference with an end-to-end solution.

**(iv)** The system is resilient against known attack models, e.g., fake fingers, replay attacks and side-channels, thereby maintaining the user's trust in the security mechanism.

## FROM THE USER'S PERSPECTIVE

As a software-only solution (see Figure 3), SonicPrint can be conveniently downloaded via publicly available platforms (e.g., Apple Store, Google Play) on the user's smart device. During installation, the user may be asked to provide permission for microphone access to record the FiSe from swipe actions. When the application is launched for the first time, the user will be asked to perform 60 swipe actions (number based on current system's capability) with different dynamics for training the underlying prediction model behind SonicPrint. Once the training process is completed, the device will be automatically locked and the user would be asked to swipe between 1~3 times for gaining access to their smart device. Considering that the signal is in the audio domain, it is preferable for the swipes to occur in close proximity to the microphone. The overall duration of setting up SonicPrint is expected to be less than 3 minutes, which is comparable to existing biometric solutions.

## ON THE TECHNOLOGICAL SPECTRUM

The foundation of SonicPrint relies on the friction-excited sonic generated by the user's fingerprint upon interaction with everyday surfaces. Although the rationale is intuitive, it introduces unique challenges when applied to the mobile computing domain:

**(1) Traceability:** FiSe is typical of low power and submerged in dynamic background noises. While this contributes to high security against stealth listeners, it can decrease the usability of FiSe if left unsolved. After FiSe has been recorded by conventional microphones, we leverage a sequence of spectral and wavelet denoising approaches [4,5] to enhance the target signal and remove the background noise. The dependency of FiSe on swiping behavior poses another challenge in tracking its position in the recorded signal. Observing the limitation of traditional thresholding techniques used in the voice domain, we employ an adaptive event detection approach using a Hidden-Markov model [6] and phase-based detection that is fine-tuned according to the roughness of the user's fingertip.

**(2) Resourcefulness:** Originating from multilevel fingerprint minutiae, FiSe also possesses enough capacity as a biometric trait. However, a statistical representation of FiSe can be influenced by the variation in the user's swiping speed or pressure during each access attempt; it is vital that the representation (in other words, features) of FiSe closely relates to the user's fingerprint. We explore a taxonomy that bridges the gap between fingerprint in the visual domain and FiSe in the audio domain by leveraging temporal, perceptual, cepstral and physical features.

**(3) Distinguishability:** SonicPrint has an advantage to be deployable in previously untouched domains (e.g., wearables, everyday materials) with a high degree of freedom in swipe actions and sensing locations. However, despite such freedom, it needs to offer high accuracy with limited to no assumption about the input FiSe signal. Therefore, we chose an ensemble classification approach with five weighted classifiers (e.g., logistic regression, linear discriminant analysis, support vector machine, random forest, Gaussian mixture model) to predict the legitimate user.

## BEYOND SMARTPHONES

For performance evaluation, we let five subjects conduct a total of 1981 swipe actions on four popular smart devices with an increasing level of curvature: Bose Headphones, Google Echo, Apple Watch Series 4 (leather strap) and Logitech mouse. The Logitech mouse comprises an inward surface while the rest are outward. The microphone is positioned near the surfaces of considered smart devices to record the FiSe during swipe action. For each device, the Equal-Error-Rate (EER) during user identification is illustrated in Figure 4. Our approach is adoptable to devices that are currently do not support any biometric technology.

## FUTURE APPLICATIONS

**Group Authentication:** Biometric technologies have transformed the user security by analyzing diverse physiological and behavioral traits via unique frameworks, e.g., multimodel, unobtrusive and continuous authentication [7]. Yet conventional biometrics provide a one-to-one connection between the measured signal and the user's identity. If users belonging to a group (e.g., family, colleagues) need to be authenticated at a single timestep (e.g., border verification in airports), multiple sensors are required with increased resolution and field-of-view. Moreover, the software algorithms need to individually assess each biometric trait making the computational time complexity similar between identifying the group together vs. each person separately. If FiSe from different groups of users can be identified without any change in system architecture, it can lead to a breakthrough in the field of mobile security research.

**Surface Identification:** Recently, object tagging without Near Field Communication (NFC) tags have received immense attention for robotics [8] and mobile applications [9]. The uniqueness of FiSe relates to the fingerprint minutiae, surface texture and the underlying composition of the human fingertip. Its dependency on surface texture raises an interesting question of whether SonicPrint can be applied for object identification.

**Gesture Recognition:** Gesture recognition has observed a significant growth in the smart environment due to its application in entertainment, gaming, motion capture and accessibility services. In particular, device-free tracking is promising, since it does not require a user to place hands/fingers attached to the device. Considering that FiSe can be acquired remotely via a microphone, SonicPrint may serve as a multi-purpose application to sense both the biometric trait and soft characteristics from a single swipe action.

## CHALLENGES AND COUNTERMEASURES

**Aging Effects:** For every material surface, a different degree of variation occurs over time. Common materials used in day-to-day activities have a tolerance to aging, but it can be accelerated under heavy load (e.g., multiple users swiping on a common device surface). In real practice, every user has their own personalized devices. Yet, it can be beneficial to explore the use of specialized materials that are more resistant to aging for superior longevity.

**Microphone Sensitivity:** SonicPrint leverages the low-cost microphone of smartphones for FiSe acquisition. Although our system shows a satisfactory performance under ideal conditions, the overall performance can be significantly improved by adopting highly sensitive microphones. These microphones can precisely detect FiSe from even swipe actions on smooth surfaces in a noisy environment. Users would not be required to perform the swipe as close to the microphone, increasing the level of freedom and user acceptance.

**Privacy:** The audible nature of FiSe makes it prone to theft via a conventional recording device. For a countermeasure, the user can be asked to perform a specialized gesture (e.g., zigzag or star pattern) during the training process. These gestures are uncommon in normal user behavior, thereby increasing the difficulty for an attacker to acquire the target FiSe outside the recognition period.

**Accuracy and Improvements:** SonicPrint achieves 84% and 98% identification rates with a single trial on a standard and high-texture smartphone surface, respectively. This is comparable to recent low-cost solutions using vibrations [10,11], gait patterns [12] and passive sensing [13] for authentication. Yet the most significant contribution of SonicPrint is its adoptability across smart devices, which is not supported by existing solutions. The proposed approach can also be used as secondary biometrics; improvements in microphone frequency response and deep learning approaches can be considered for our future exploration. ■

**Aditya Singh Rathore** is a fourth-year PhD candidate in the Computer Science and Engineering Department at the University at Buffalo, SUNY, where he also received his BS degree in Computer Engineering in 2017. His research focuses on mobile security, biometrics and human-computer interaction.

**Chenhan Xu** is a third-year PhD student in the Computer Science and Engineering Department in the University at Buffalo, SUNY. His current research interests include Internet of Things, mobile computing, human-computer interaction, and mobile health.

**Wenyao Xu** is an Associate Professor of Computer Science and Engineering Department in the University at Buffalo, SUNY. His research area includes mobile sensing, mobile health and mobile security. He received his PhD degree from the University of California, Los Angeles.

## REFERENCES

[1] A. Akay. 2002. Acoustics of friction, *The Journal of the Acoustical Society of America*, vol. 111, no. 4, 1525–1548.

[2] Raffaele Cappelli and Matteo Ferrara. 2012. A fingerprint retrieval system based on level-1 and level-2 features. *Expert Systems with Applications*, 39.12: 10465-10478.

[3] Dalibor Mitrović, Matthias Zeppelzauer, and Christian Breiteneder. 2010. Features for content-based audio retrieval. *Advances in Computers*. Vol. 78. Elsevier. 71-150.

[4] S. Kamath and P. Loizou. 2002. A multi-band spectral subtraction method for enhancing speech corrupted by colored noise, *ICASSP*, vol. 4. *Citeseer*, 44 164–44 164.

[5] D.B. Percival and A.T. Walden, Wavelet Methods for Time Series Analysis. *Cambridge University Press*, 2006, vol. 4.

[6] Y. Bi, M. Lv, C. Song, W. Xu, N. Guan, and W. Yi. 2015. Autodietary: A wearable acoustic sensor system for food intake recognition in daily life, *IEEE Sensors Journal*, vol. 16, no. 3, 806–816.

[7] M. Abuhamad, A. Abusnaina, D. Nyang, and D. Mohaisen. 2020. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A survey. *arXiv,* 2001.08578.

[8] W. Yuan, S. Dong, and E. H. Adelson. 2017. Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, vol. 17, no. 12, 2762.

[9] K. Ali and A. X. Liu. 2020. Fine-grained vibration based sensing using a smartphone. *arXiv,* 2001.08578.

[10] J. Liu, C. Wang, Y. Chen, and N. 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*. ACM, 73–87.

[11] J. Li, K. Fawaz, and Y. Kim, Velody: Nonlinear vibration challenge-response for resilient user authentication, in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security.* ACM, pp. 1201–1213.

[12] Y. Ren, Y. Chen, M.C. Chuah, and J. Yang. 2013. Smartphone based user verification leveraging gait recognition for mobile healthcare systems, in *IEEE International Conference on Sensing, Communications and Networking (SECON).* IEEE, 149–157.

[13] W.-H. Lee and R.B. Lee, Multi-sensor authentication to improve smartphone security. 2015. *Proceedings of the International Conference on Information Systems Security and Privacy (ICISSP)*. IEEE, 1–11.