PrintListener: Uncovering the Vulnerability of Fingerprint Authentication via the Finger Friction Sound

Man Zhou, Shuao Su, QianWang, Qi Li, Yuting Zhou,
Xiaojing Ma, Zhengxiong Li

NDSS 2024

Outline

- Motivation
- PrintListener
- Attack Evaluation
- Conclusion

Finger-swiping During Audio/video Calls

Skype

Google Meet

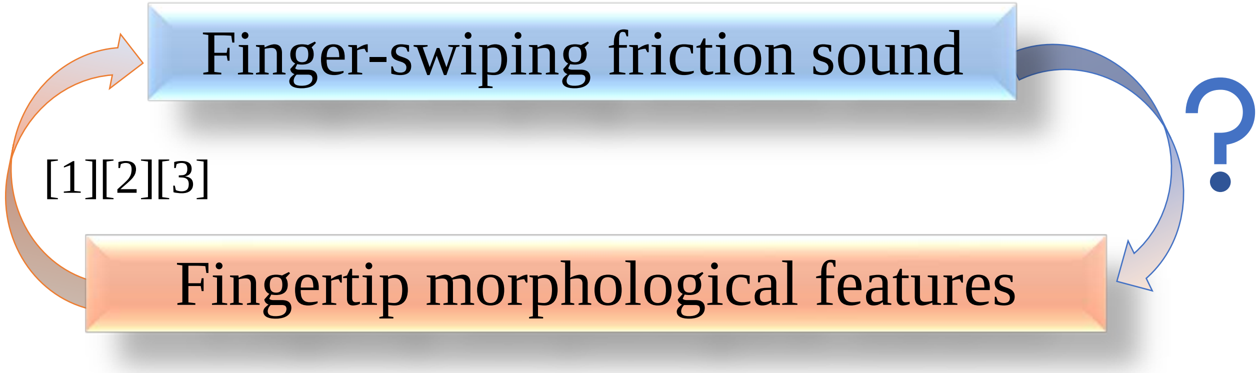
Zoom

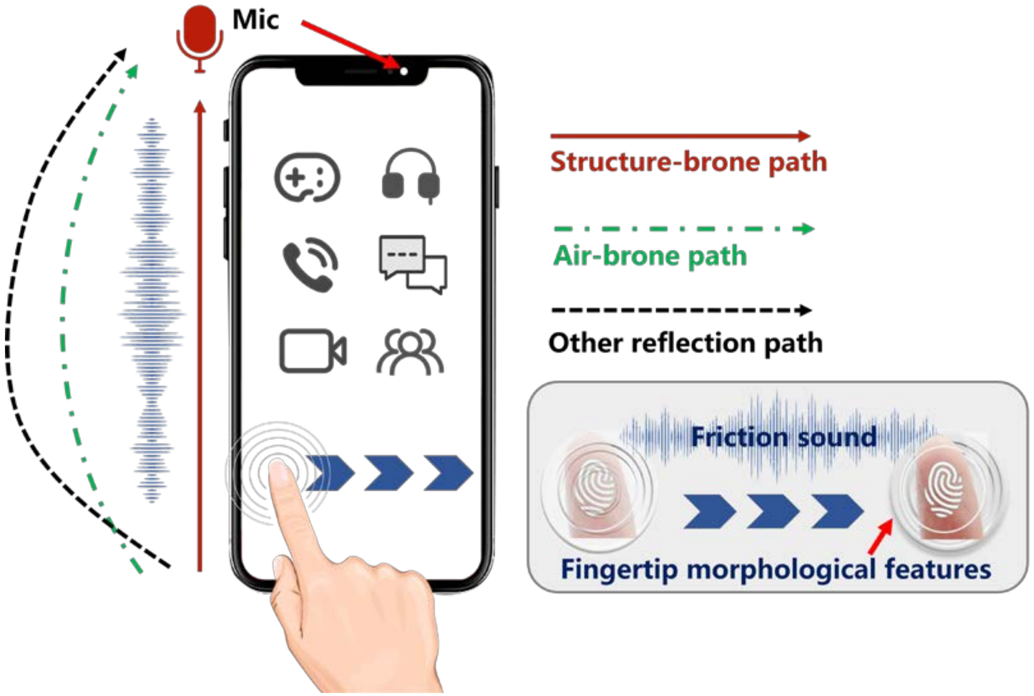
WeChat

Discord

Malware

Acoustic Principle



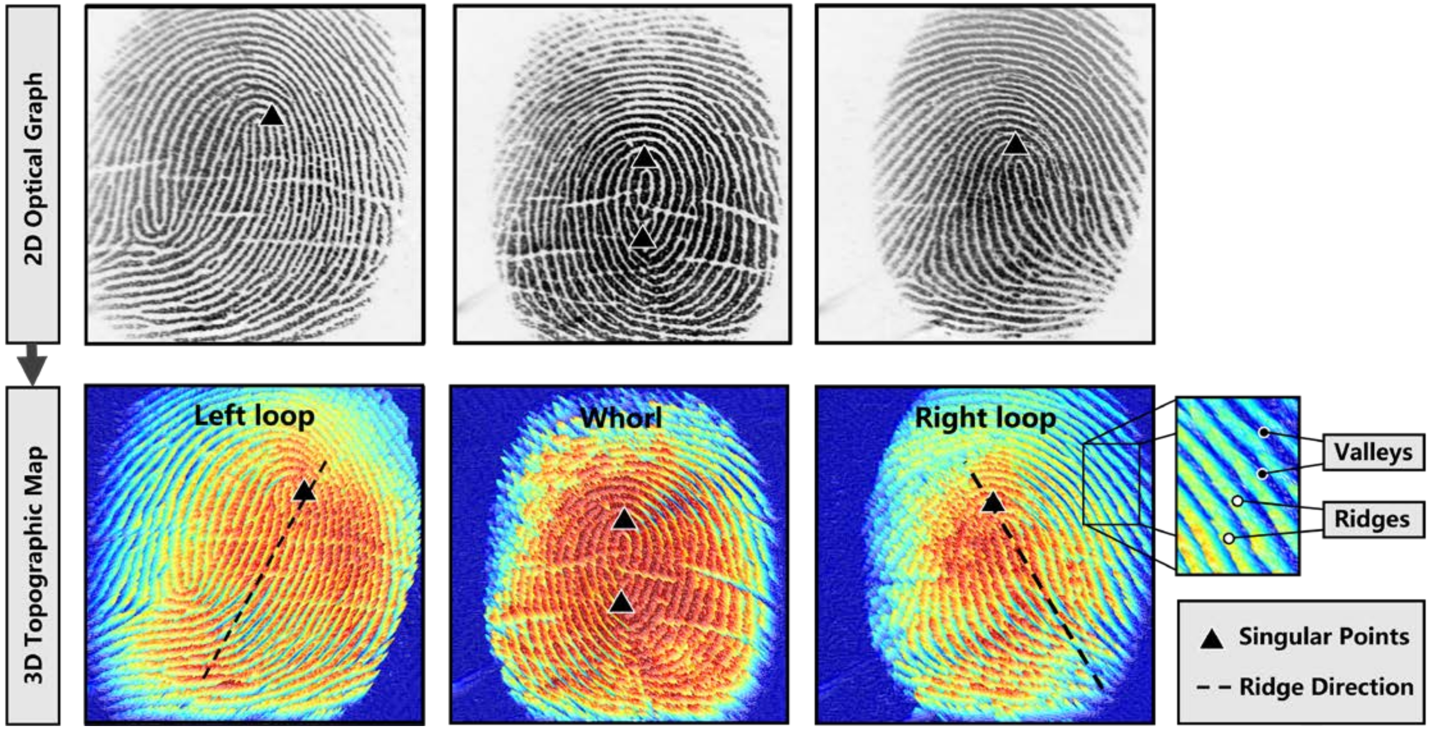


- The production of finger-swiping sound
  - Friction (the elastic deformation between the fingertips and the screen)
  - Dynamics (the vibrations and waves propagate between the finger and the screen)
  - Acoustics (audible roughness sound radiates from the finger to the microphone)

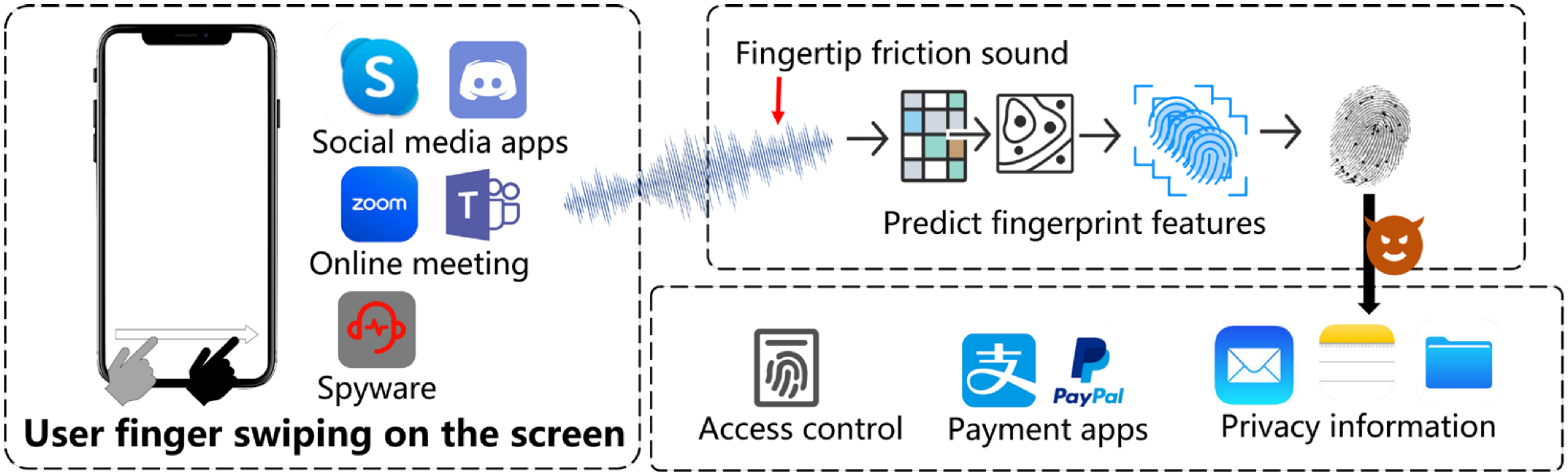
[1]A. S. Rathore, W. Zhu, A. Daiyan, C. Xu, K. Wang, F. Lin, K. Ren, and W. Xu, "Sonicprint: a generally adoptable and secure fingerprint biometrics in smart devices," in Proc. of ACM MobiSys, 2020, pp. 121–134.
[2] Z. Shu, Z. Wang, G. Yang, C. Zang, Z. Ma, F. Lin, and K. Ren, "Fingersound: A low-cost and deployable authentication system with fingertip sliding sound," in Proc. of IEEE ICPADS, 2023, pp. 33–40.
[3] M. Zhou, Y . Zhou, S. Su, Q. Wang, Q. Li, S. Hu, C. Y u, and Z. Li, "Fingerpattern: Securing pattern lock via fingerprint-dependent friction sound," IEEE Transactions on Mobile Computing, 2023.

Acoustic Principle

- The ridges of fingerprints reduce the contact area between the finger pad and the screen, resulting in variations in frictional radiation of air, solid vibration, and wave propagation modes

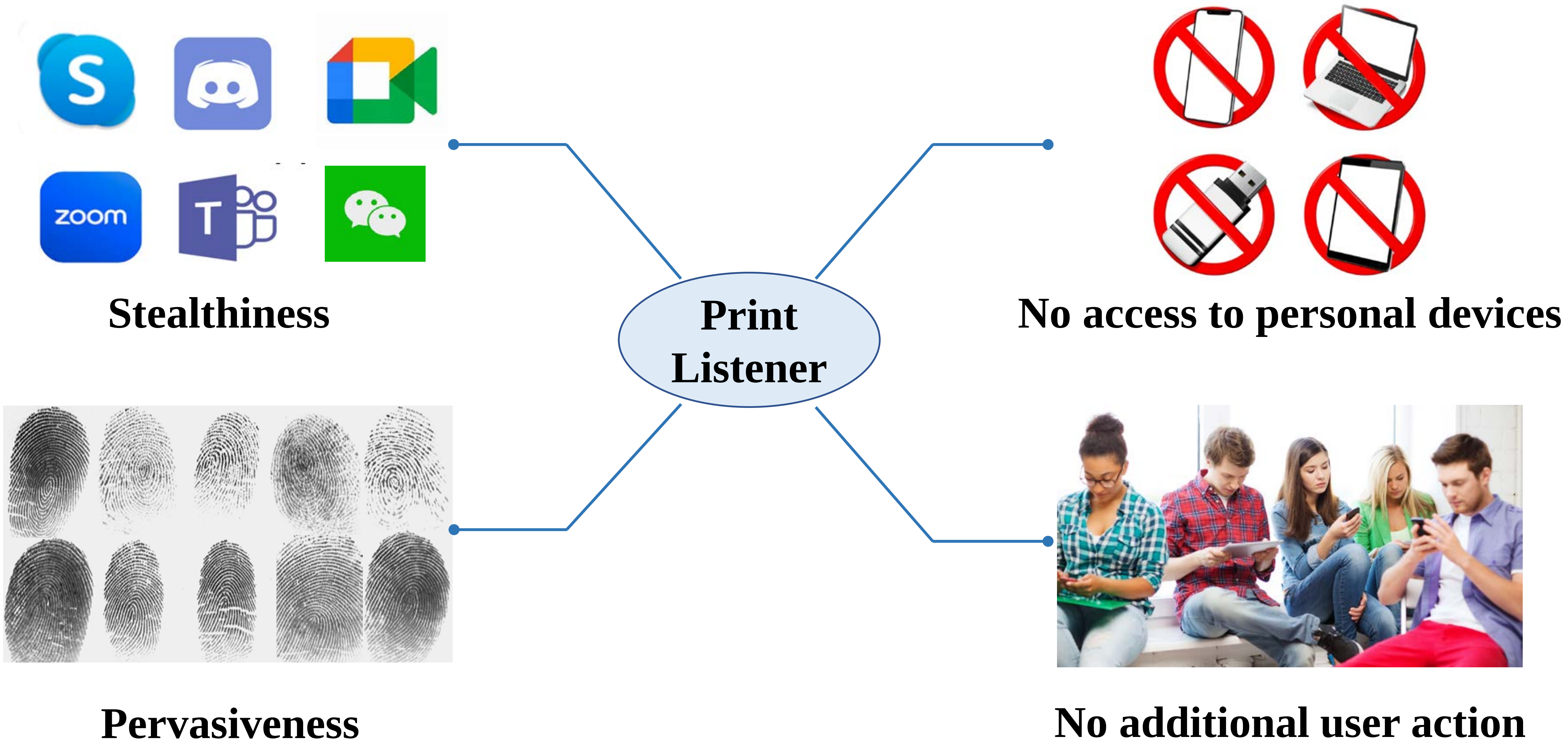


Our Idea



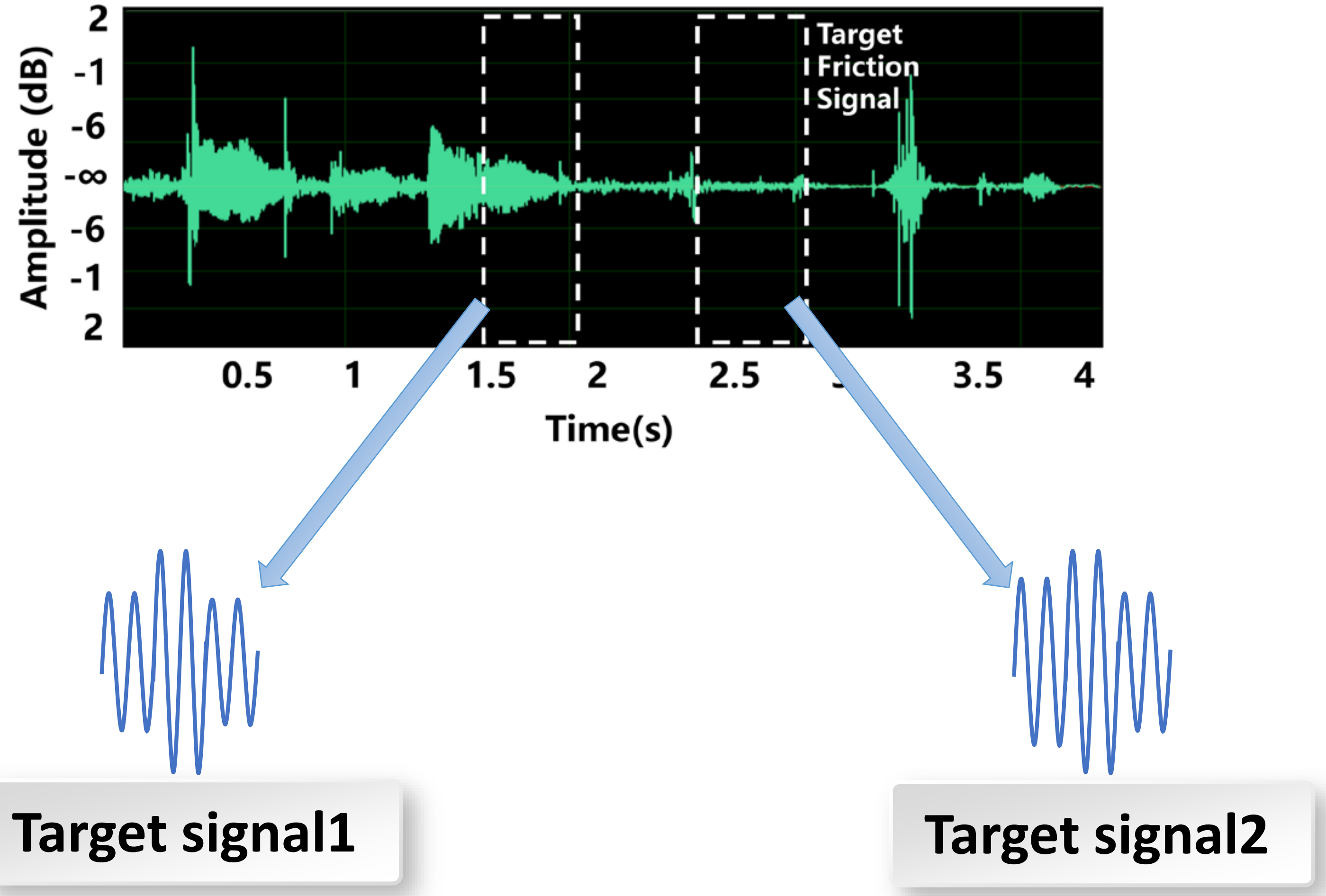
- We propose PrintListener, a novel attack to predict fingerprint patterns by leveraging users' swiping actions and then synthesize a stronger fingerprint minutiae attack templates

Advantages of PrintListener



Challenges

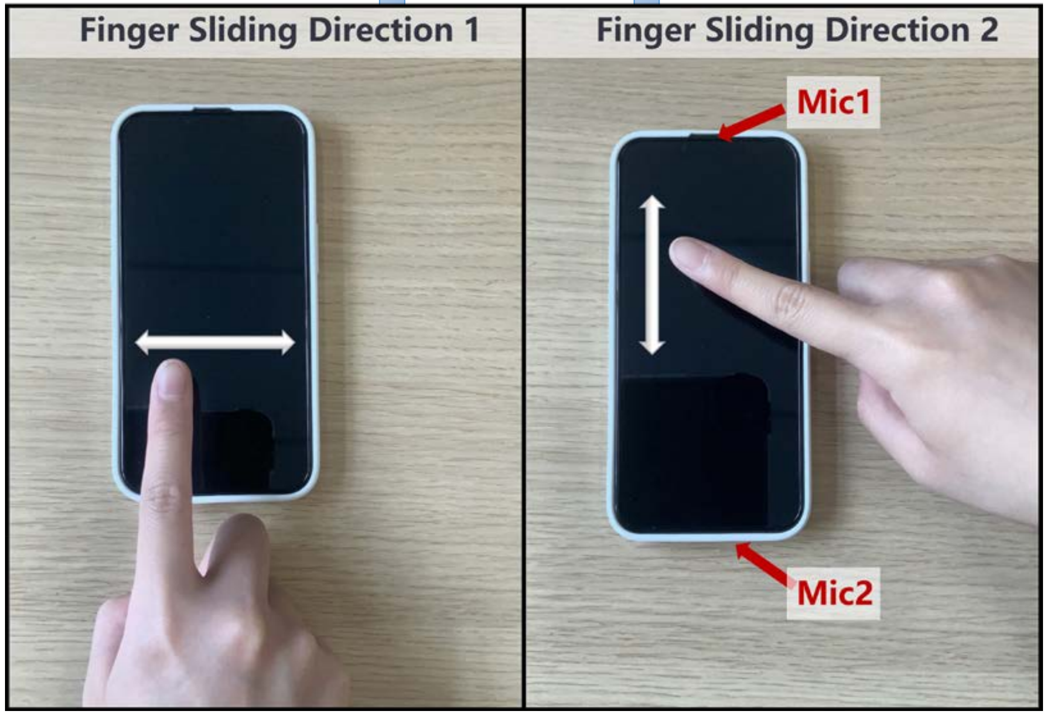
- The sound intensity of finger friction from users is extremely weak, typically ranging from 0.2 to 0.8 S



Challenges

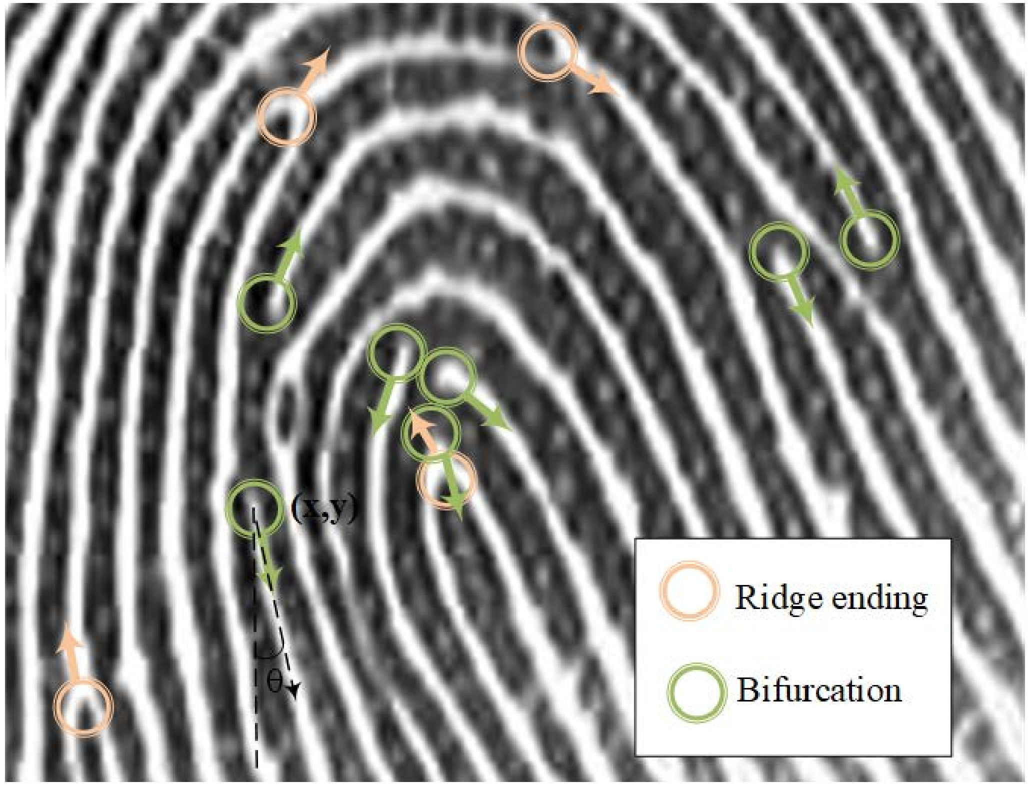
- Friction sound characteristics are often influenced by users' physiological and behavioral features

pressure **velocity** **varied swiping trajectory**



Challenges

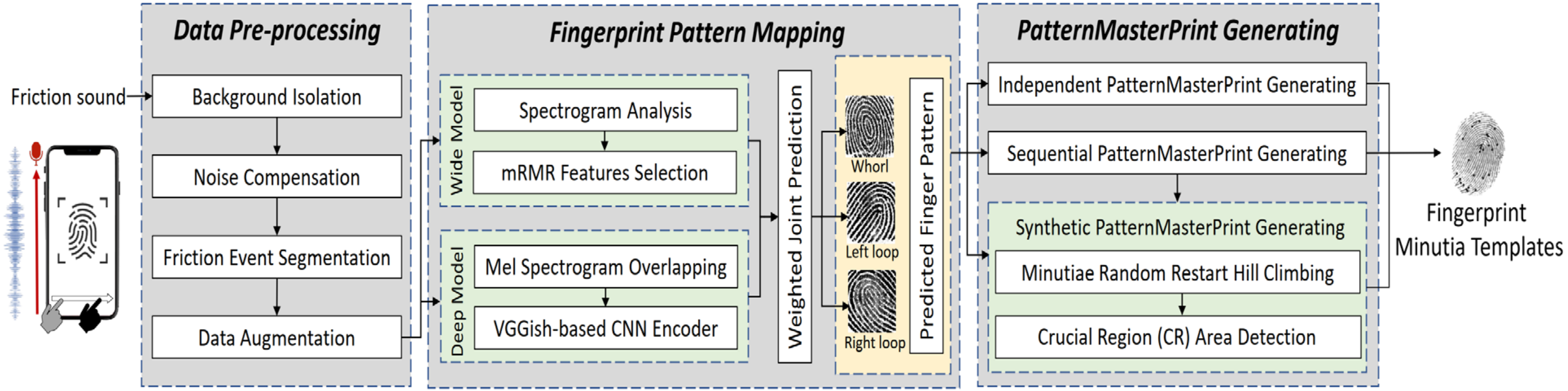
- After inferring the primary pattern features of fingerprints, the potential search space for the secondary features corresponding to fingerprints of the same pattern is vast



Outline

- Motivation
- PrintListener
- Attack Evaluation
- Conclusion

System Overview



Data Pre-processing

☐ BackGround Noise Isolation

- A finite impulse response high-pass filter (FIR) with a 4 kHz passband to eliminate low-frequency noise while preserving the fingerprint information

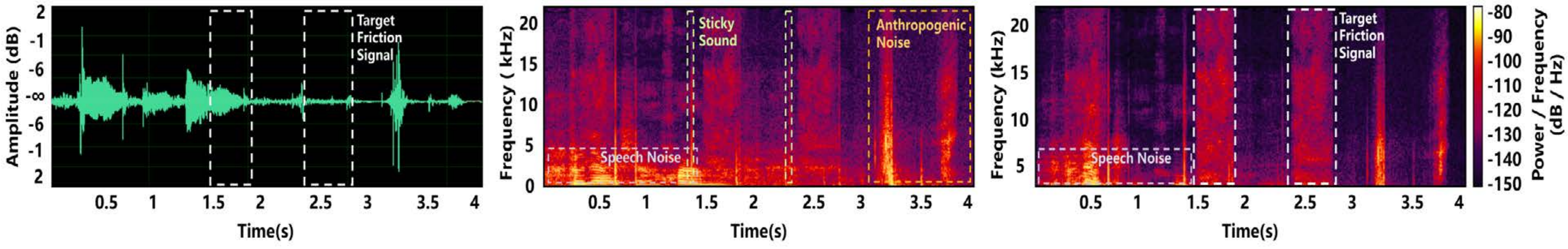
Data Pre-processing

☐ BackGround Noise Isolation

- A finite impulse response high-pass filter (FIR) with a 4 kHz passband to eliminate low-frequency noise while preserving the fingerprint information

☐ Noise Compensation

- To enhance the target signal degraded by additive noise without introducing any distortion

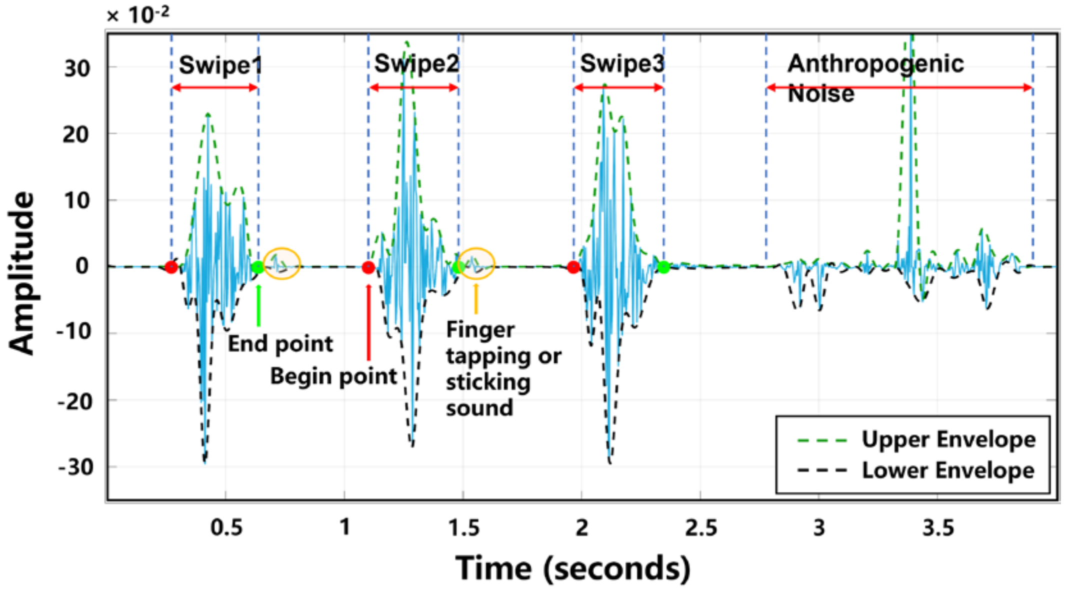


(a) Oscillogram of original friction sound signal (b) Spectrogram of original friction sound signal (c) Spectrogram of Noise-compensated friction sound signal

Data Pre-processing

☐ Friction Event Segmentation

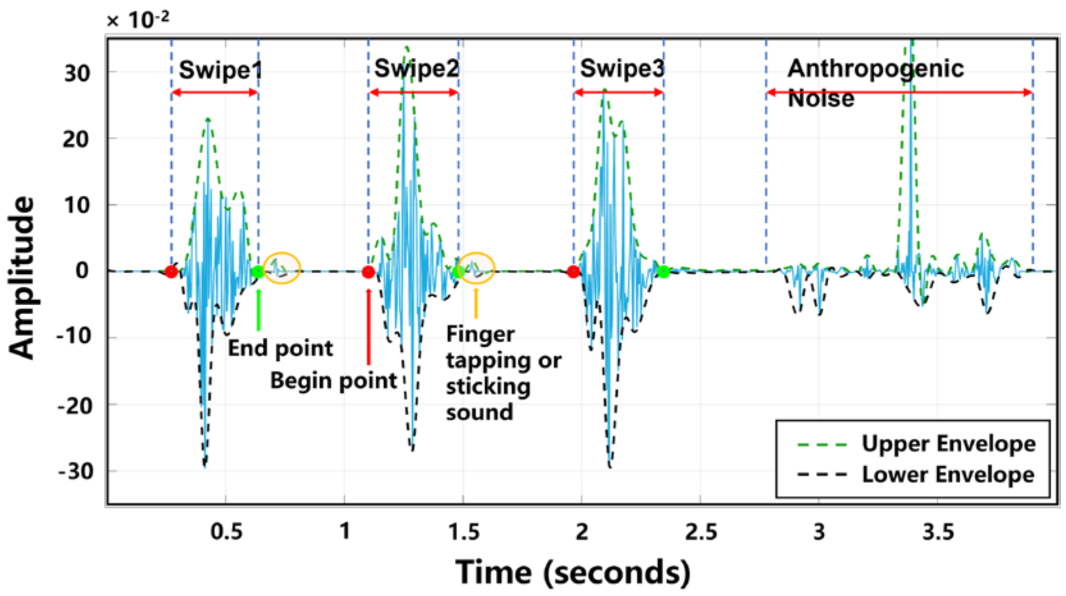
- Step1: Silent regions exclusion
- Step2: Full-frequency energy verification
- Step3: Duration verification



Data Pre-processing

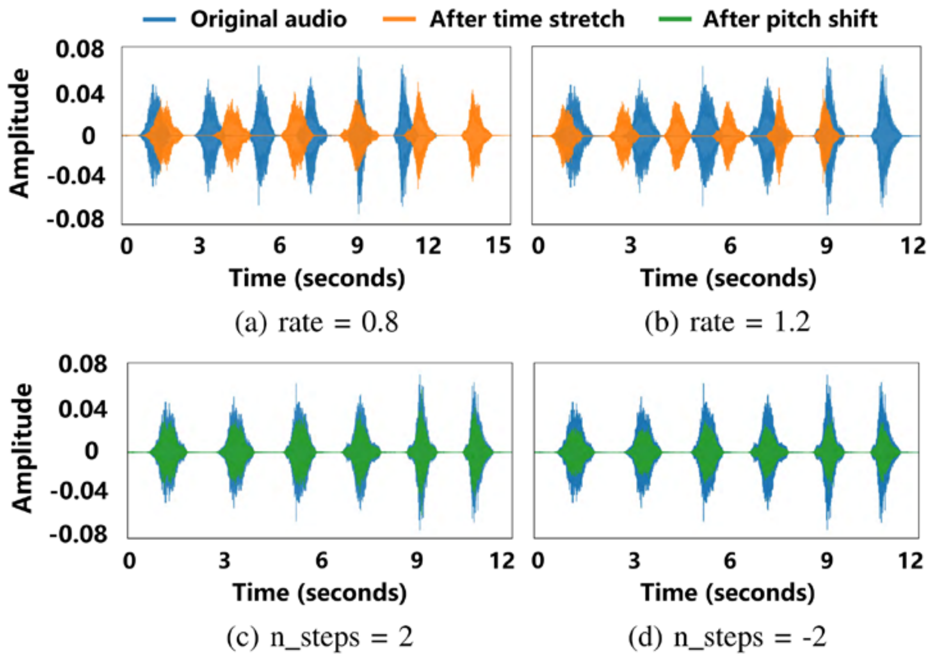
☐ Friction Event Segmentation

- Step1: Silent regions exclusion
- Step2: Full-frequency energy verification
- Step3: Duration verification



☐ Data Augmentation

- Time Stretch
- Pitch Shift



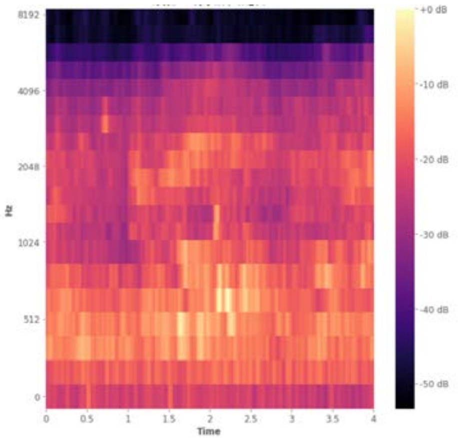
Fingerprint Pattern Mapping

◻ Interpretable Audio Features Extraction

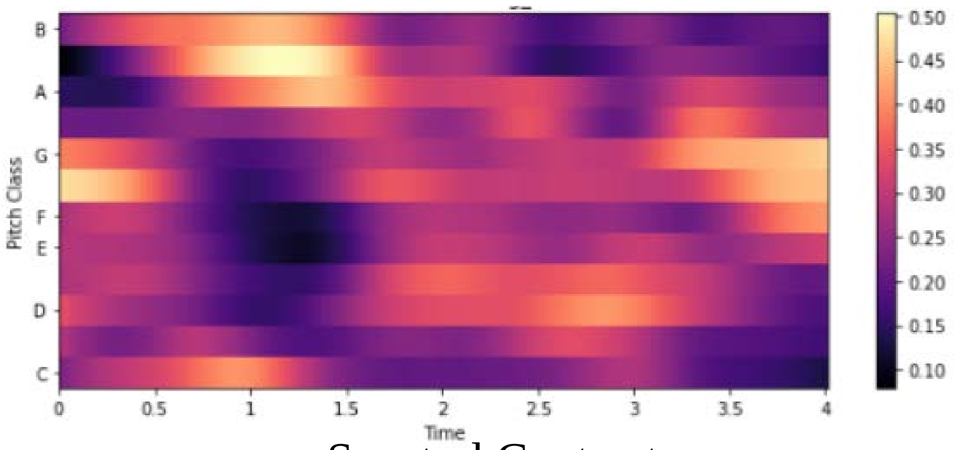
- Identifying a candidate feature set (6 frequency-domain features and 3 cepstral-domain features)

TABLE II: Selected interpretable audio features.

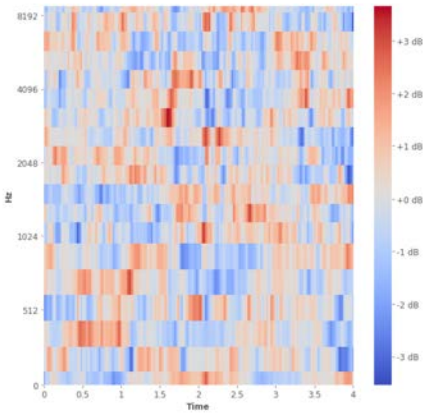
Domain	Feature	Feature vector
Frequency	LSF	$f1 - f12$
	Chroma	$f13 - f24$
	Spectral Kurtosis	$f25$
	Spectral Skewness	$f26$
	Spectral Contrast	$f27 - f33$
	Spectral Centroid	$f34$
Cepstral	MFCC	$f35 - f73$
	LPCC	$f74 - f86$
	RASTA-PLP	$f87 - f99$



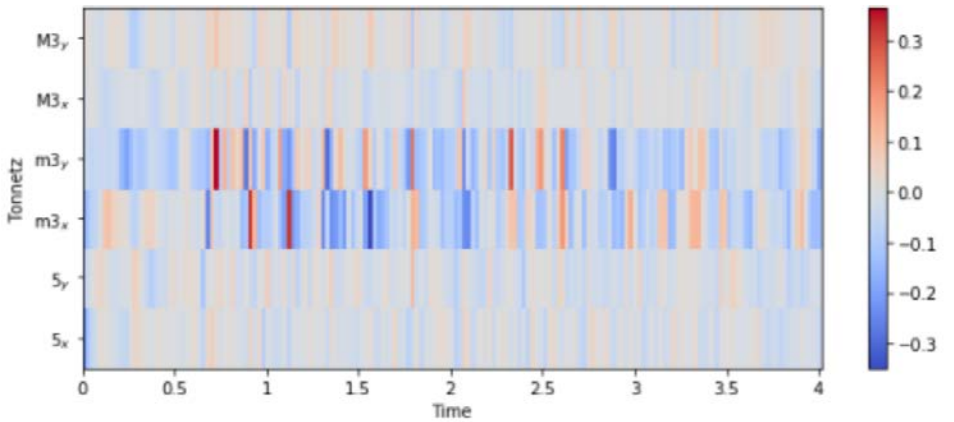
Mel spectrogram



Spectral Contrast



MFCC



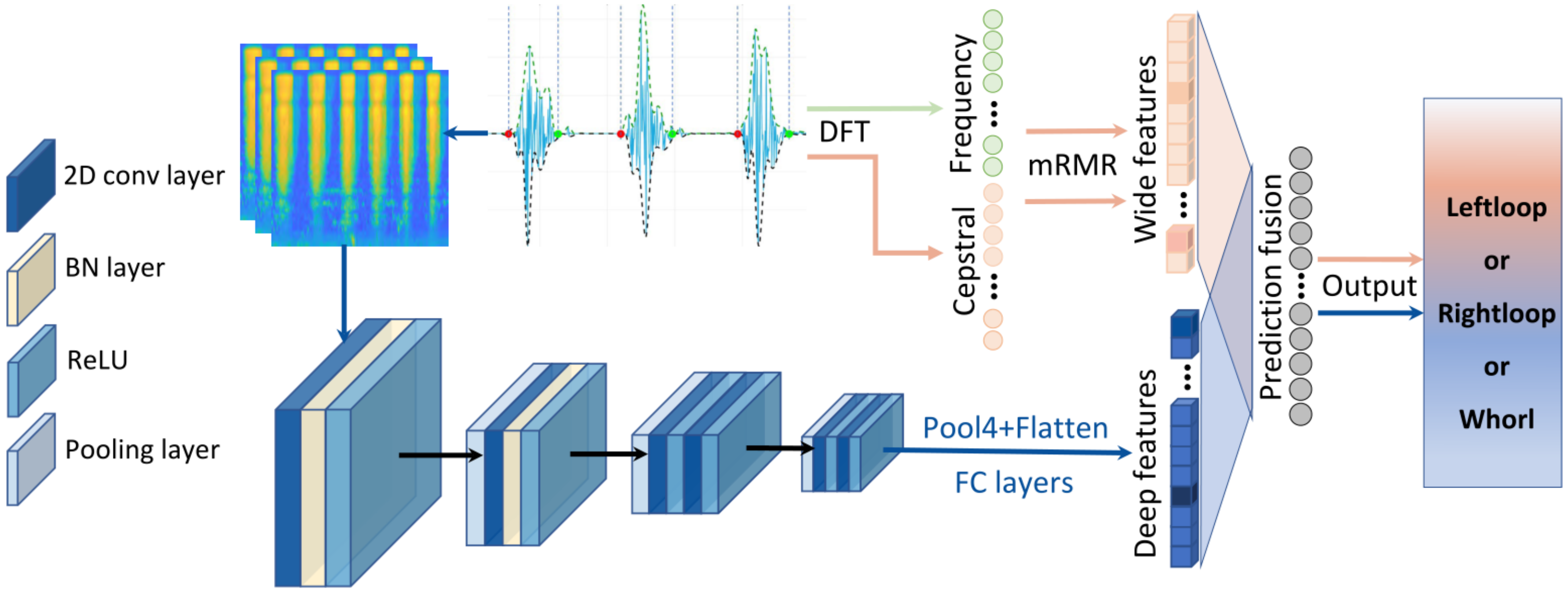
Spectral Centroid

Fingerprint Pattern Mapping

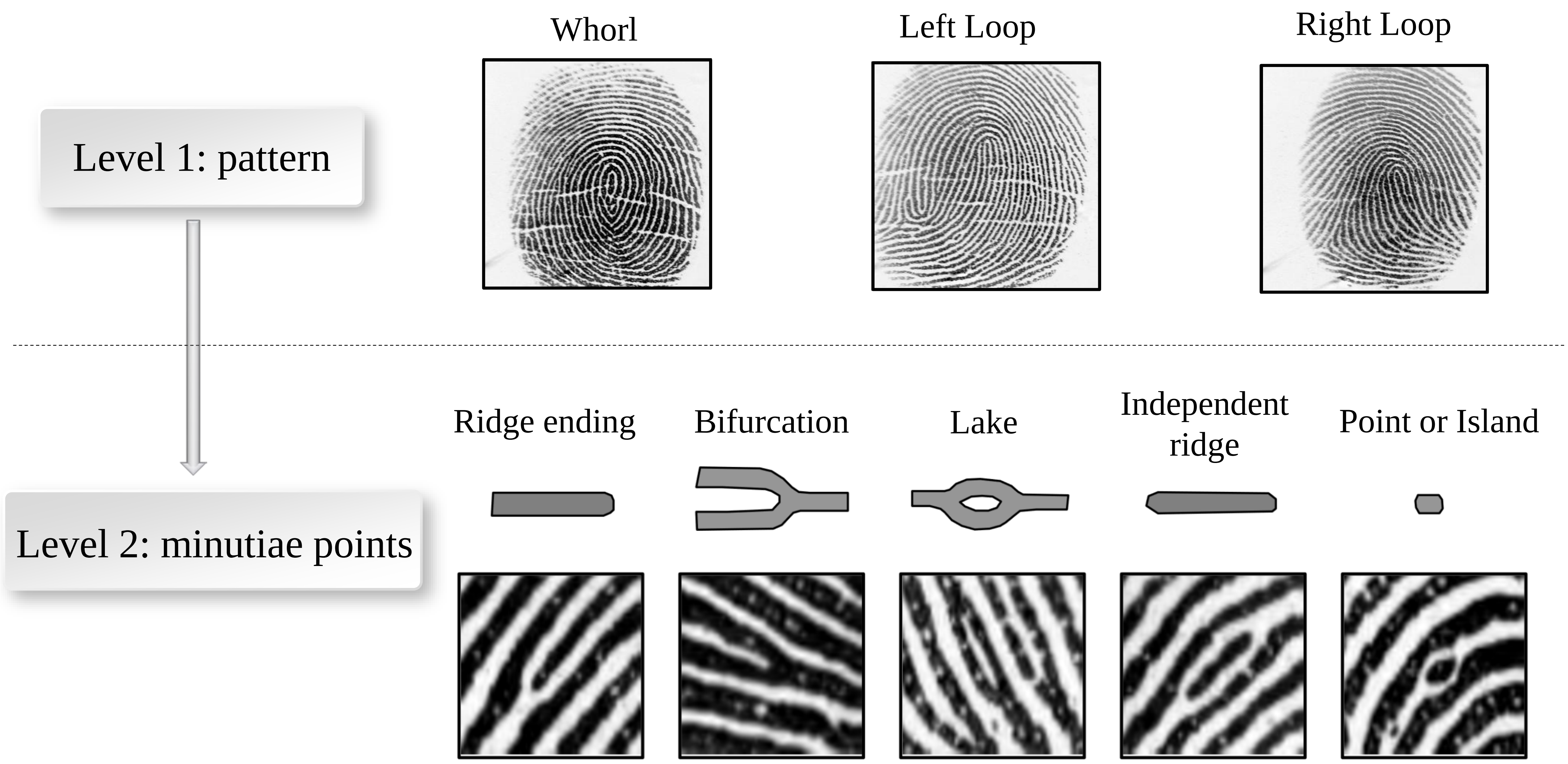
☐ Deep Representation Features Extraction

- Learning representative acoustic features using a pretrained VGGish-based CNN Encoder

☐ Weighted Joint Prediction



PatternMasterPrint Generating



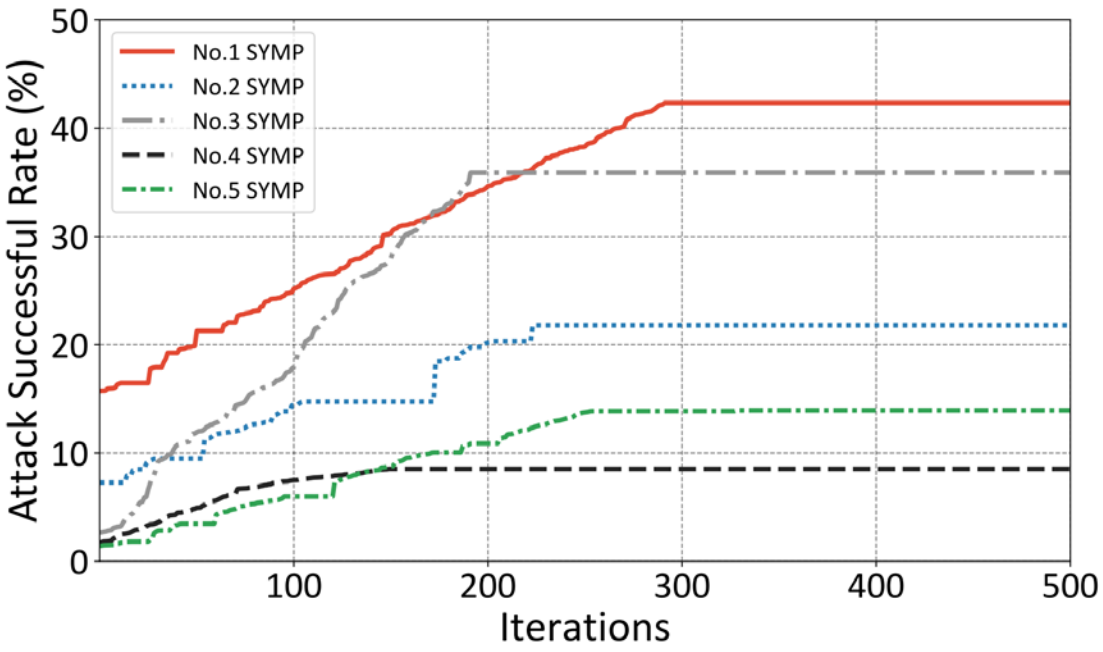
PatternMasterPrint Generating

☐ Crucial Region Area Detection

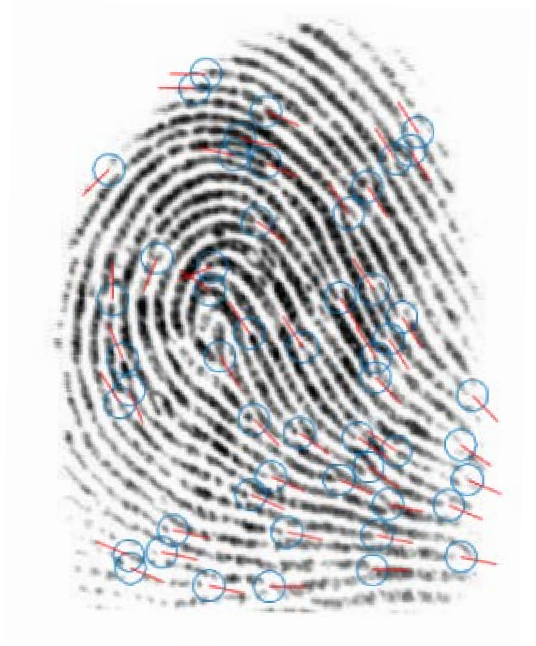
- Find the area with a high probability of fingerprint minutiae collision in the fingerprint image

☐ Minutiae Random Restart Hill Climbing

- Throughout the iterative process, the best performing detail template serves as the stored state



Synthetic PatternMasterPrint



Genuine Fingerprint

Outline

- Motivation
- PrintListener
- Attack Evaluation
- Conclusion

Evaluation Setup

☐ Dataset

- 65 subjects in the data collection (24 females and 41 males)
- Compiled friction sound datasets under different devices, e.g., Pixel 4, iPhone 13 and Samsung A20S, different experiment environments, e.g., conference room, office and playground
- Compiled fingerprint datasets of PatternFinger(the complete fingerprint dataset), FingerPassDB7 (the partial fingerprint dataset) and Livedet2011 ItalData (the complete fingerprint dataset)

☐ Metrics

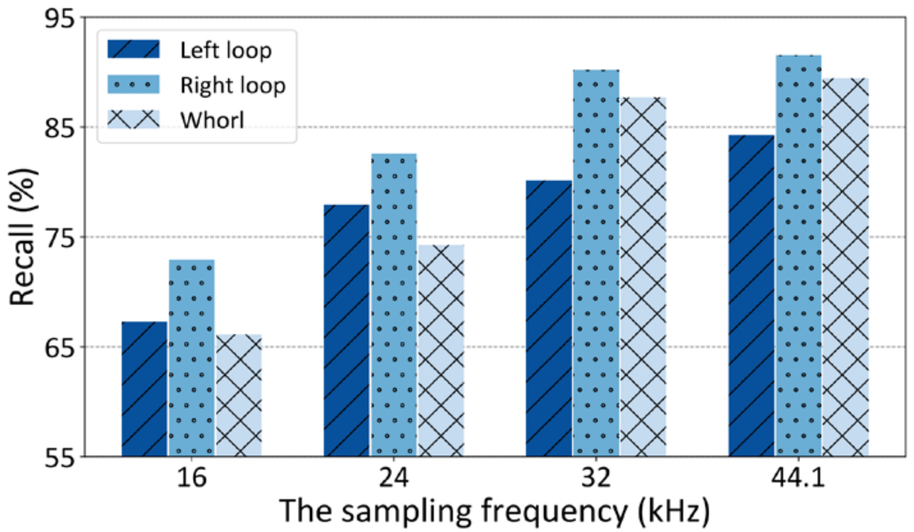
- Weighted-average precision (wP), Weighted-average recall (wR), F1 score
- Weighted attack success rate (wASR)

Results by Using Different Classifiers

Module	Accuracy	P (left loop/right loop/whorl)	R (left loop/right loop/whorl)	F_1 score
YAMnet+KNN	0.709	0.760 / 0.712 / 0.657	0.747 / 0.694 / 0.685	0.709
YAMnet+Decision Tree	0.820	0.828 / 0.758 / 0.881	0.840 / 0.804 / 0.816	0.821
YAMnet+Ramdom Forest	0.731	0.718 / 0.767 / 0.705	0.614 / 0.825 / 0.755	0.731
YAMnet+Adaboost	0.776	0.778 / 0.751 / 0.797	0.762 / 0.715 / 0.851	0.775
VGGish-like+KNN	**0.884**	**0.939 / 0.865 / 0.857**	**0.915 / 0.895 / 0.887**	**0.886**
VGGish-like+Decision Tree	0.766	0.777 / 0.791 / 0.736	0.800 / 0.672 / 0.825	0.767
VGGish-like+Ramdom Forest	0.739	0.711 / 0.779 / 0.725	0.752 / 0.835 / 0.632	0.739
VGGish-like+Adaboost	0.774	0.831 / 0.737 / 0.764	0.696 / 0.736 / 0.889	0.776
Resnet34+KNN	0.746	0.739 / 0.706 / 0.815	0.763 / 0.841 / 0.633	0.750
Resnet34+Decision Tree	0.686	0.746 / 0.673 / 0.652	0.644 / 0.697 / 0.718	0.688
Resnet34+Ramdom Forest	0.753	0.795 / 0.735 / 0.735	0.712 / 0.681 / 0.865	0.754
Resnet34+Adaboost	0.686	0.647 / 0.900 / 0.658	0.771 / 0.612 / 0.675	0.710

- **VGGish-like+KNN** outperforms the other networks with an accuracy of 88.4%

Impact of Sampling Rate

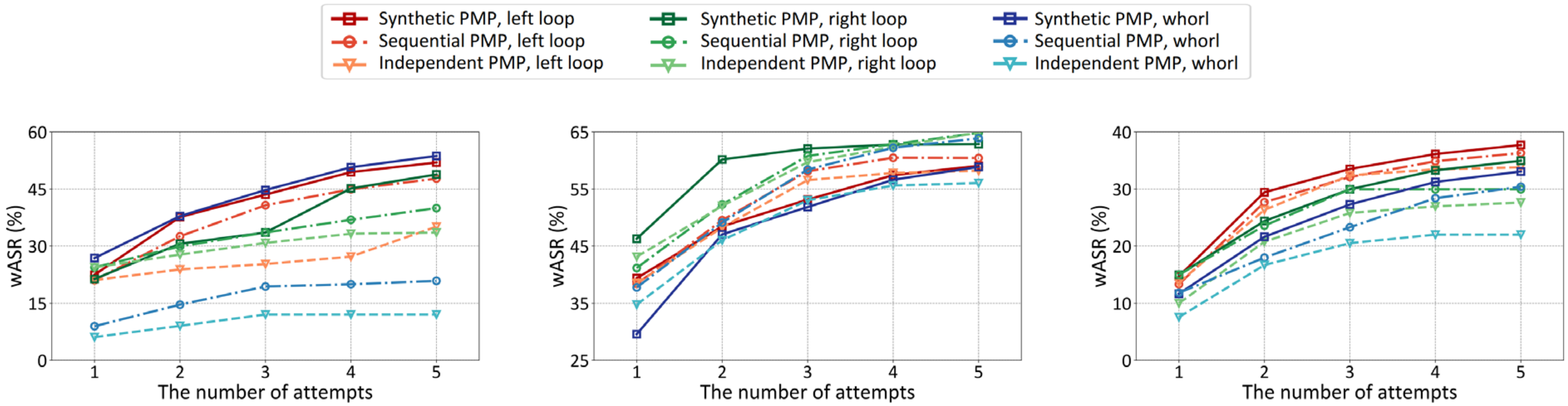


Apps	Sampling rates (kHz)
Skype	8 / 12 / 16 / 24
FaceTime	8 / 12 / 16 / 24
Google Meet	24 / 32
Microsoft Teams	16 / 32
Wecom	16 / 24

- The recall in classifying the fingerprint pattern gradually decreases as the sampling rate decreases
- 32 kHz is a commonly used sampling rate in most audio and video social networking software

Impact of Fingerprint Integrity

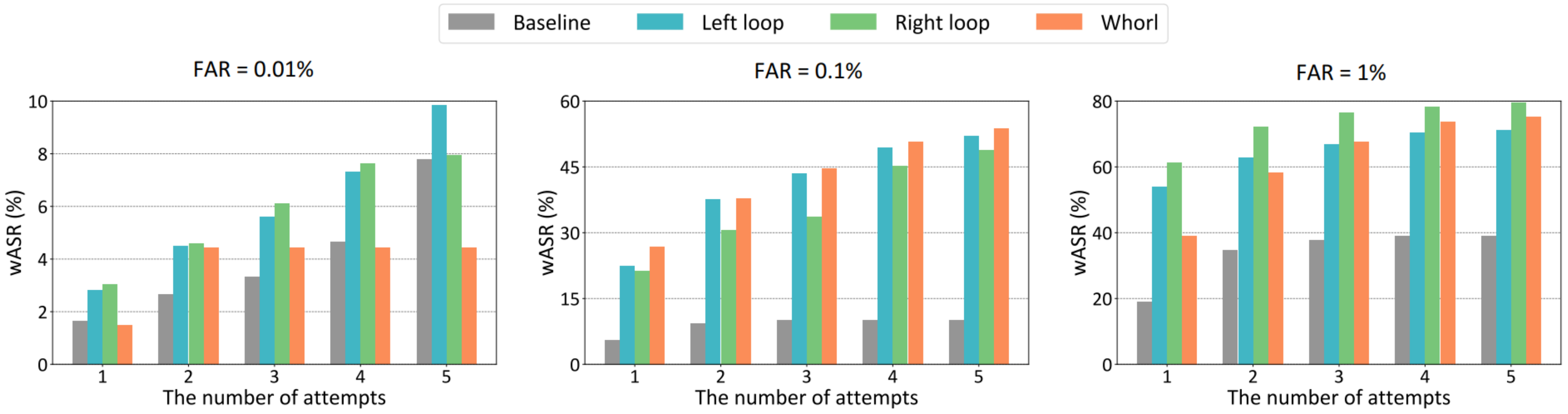
- The wASR of partial fingerprints is significantly higher than that of complete fingerprints
- The wASR of synthetic PatternMasterPrints is generally higher than that of sequential PatternMasterPrints and independent PatternMasterPrints



(a) PatternFinger (b) FingerPassDB7 (c) Livedet2011 ItalData

Impact of FAR Security Setting

- The wASR decreases at a lower FAR value setting (higher security), while more test subjects can be successfully attacked at a higher FAR value setting (lower security)

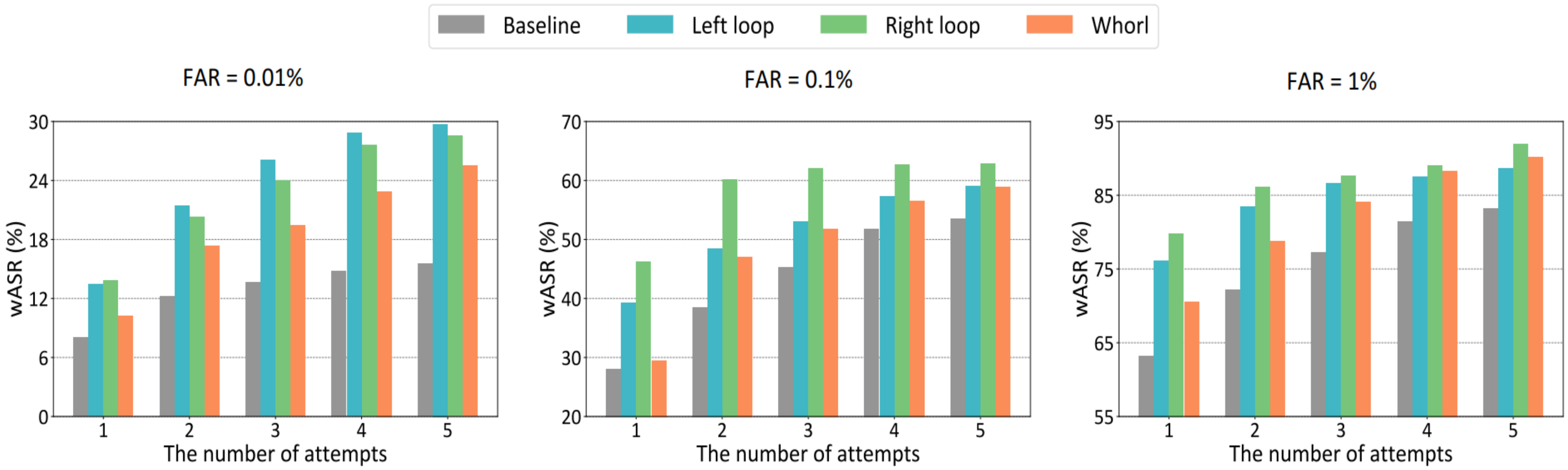


(a) PatternFinger

- The attack success rates are **52%**, **48.8%**, and **53.7%** of users with the left loop, right loop, and whorl fingerprints within 5 attempts while FAR=0.1%

Baseline Comparisons

- The MasterPrint sequences selected through pattern prediction generally have higher attack success rates than those without pattern prediction

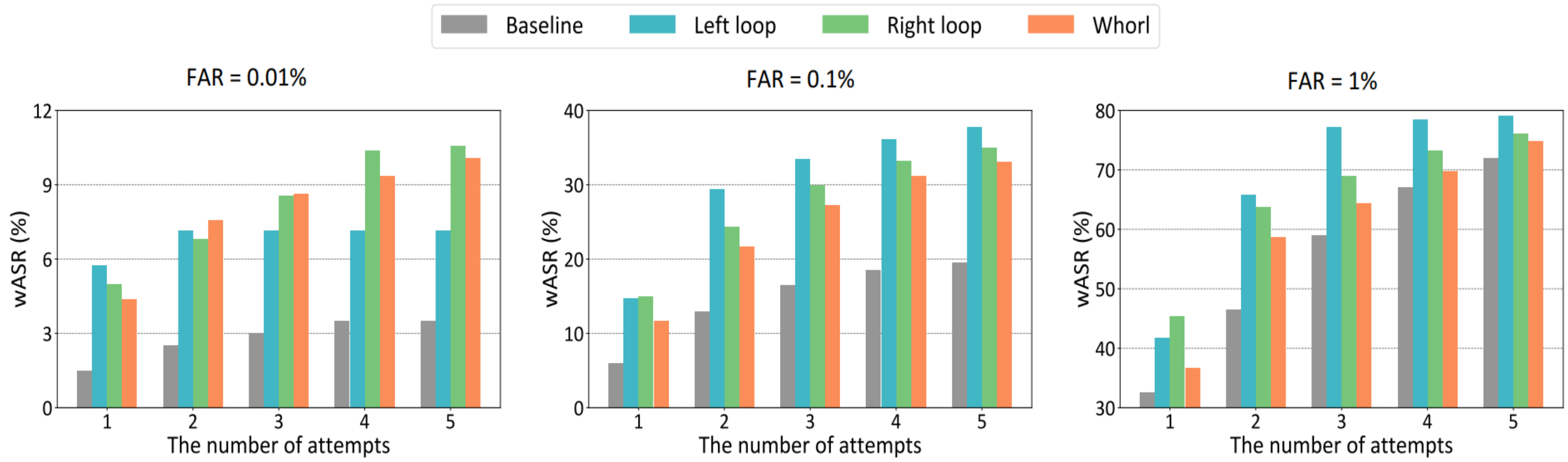


(b) FingerPassDB7

- Partial Fingerprints
  - At the highest security FAR setting of 0.01%, PrintListener achieves the average wASR of **27.9%** within 5 attempts

Baseline Comparisons

- The MasterPrint sequences selected through pattern prediction generally have higher attack success rates than those without pattern prediction



(c) Livedet2011 ItalData

- Complete Fingerprints
  - At the highest security FAR setting of 0.01%, partial achieves the average wASR of **9.3%** within 5 attempts

Outline

- Motivation
- PrintListener
- Attack Evaluation
- Conclusion

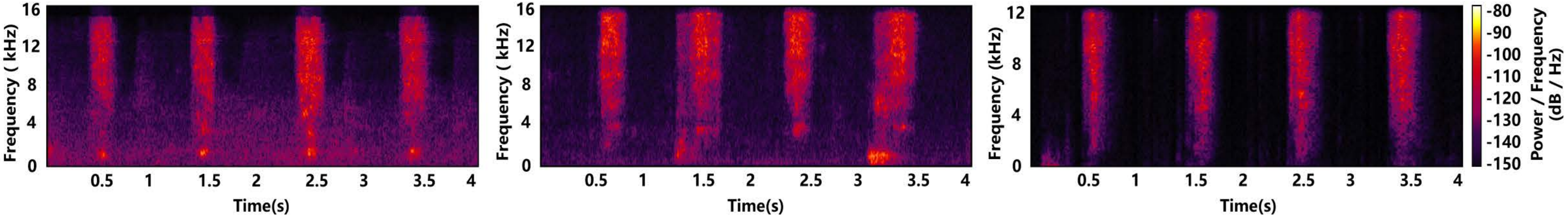
Defense

- **Correct some users' habit**
  - Avoid performing swiping operations during call

- **Audio/video social and communication apps**
  - Lower audio sample rates
  - Destroy finger frictional sound features
  - Implement pop-up reminders

Discussion

□ Attack feasibility via social networking apps



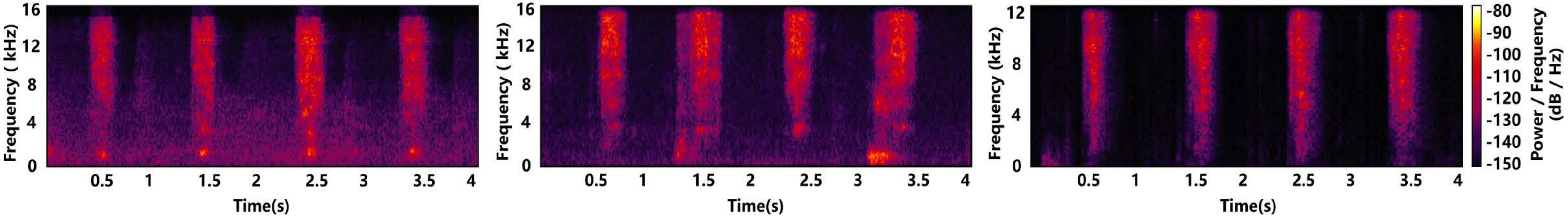
(a) Google Meet sound signal

(b) Microsoft teams sound signal

(c) Wecom sound signal

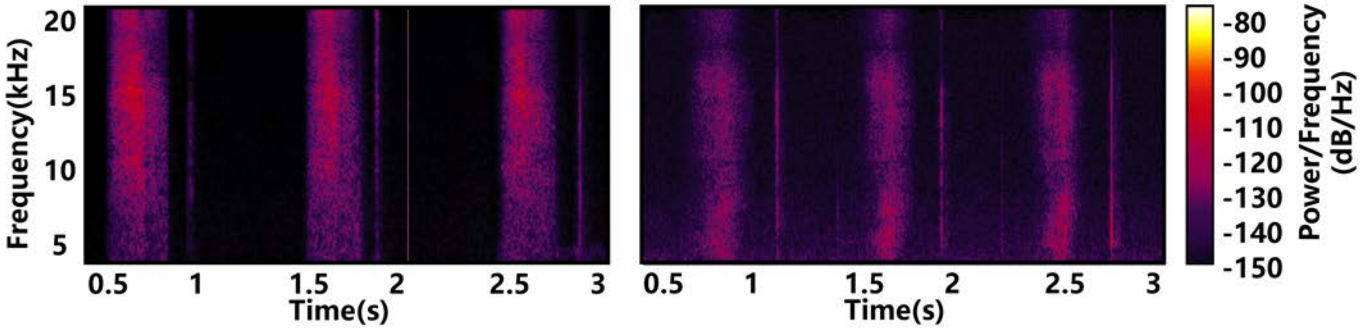
Discussion

☐ Attack feasibility via social networking apps



(a) Google Meet sound signal (b) Microsoft teams sound signal (c) Wecom sound signal

☐ Frictional sound on different films



(a) Matte film (b) Glossy film

Summary

- We uncover a new side-channel attack on fingerprint and propose PrintListener, which leverages users' swiping actions on the screen to identify the fingerprint pattern and conduct more powerful dictionary attacks

- PrintListener can automatically capture the pattern features of fingerprints from a large number of raw recordings and generate targeted synthetic PatternMasterPrints

- Extensive experimental results in real-world scenarios show that Printlistener has strong attack power on fingerprint authentication

Thank you!

Q & A