BiLock: User Authentication via Dental Occlusion Biometrics

Yongpan Zou[#], Meng Zhao[#], Zimu Zhou[*], Jiawei Lin[#], Mo Li[$], Kaishun Wu[#]

[#]*College of Computer Science & Software Engineering, Shenhzhen Univ.*

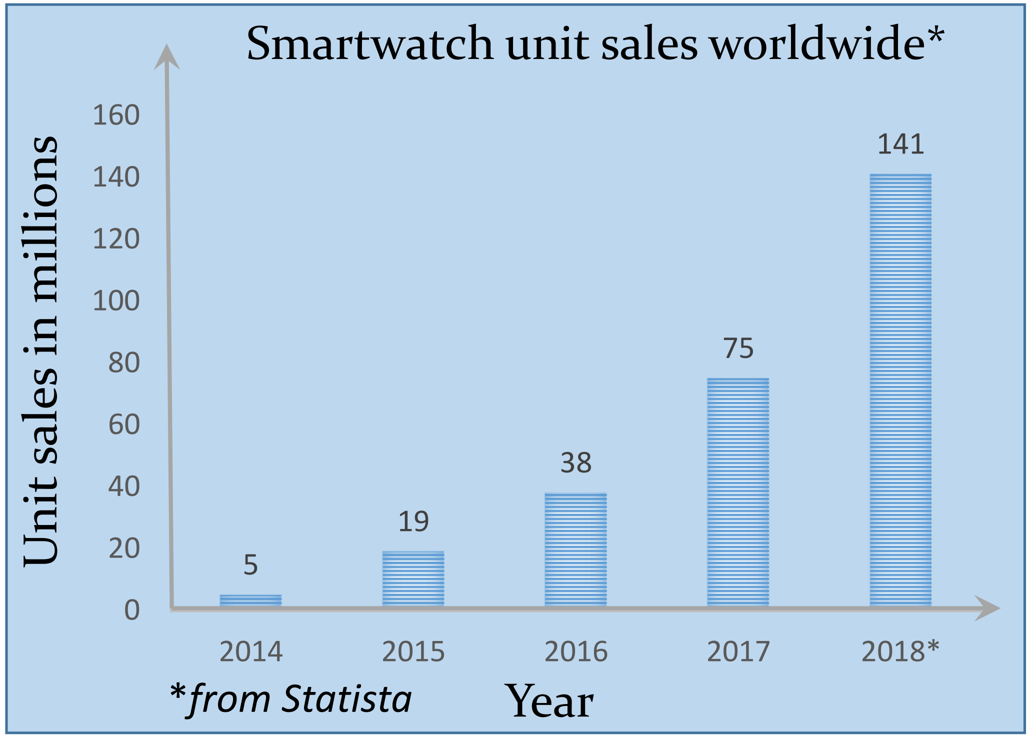
[*]*Computer Engineering and Networks Laboratory, ETH Zurich*

[$]*School of Computer Science and Engineering, Nanyang Technological University*

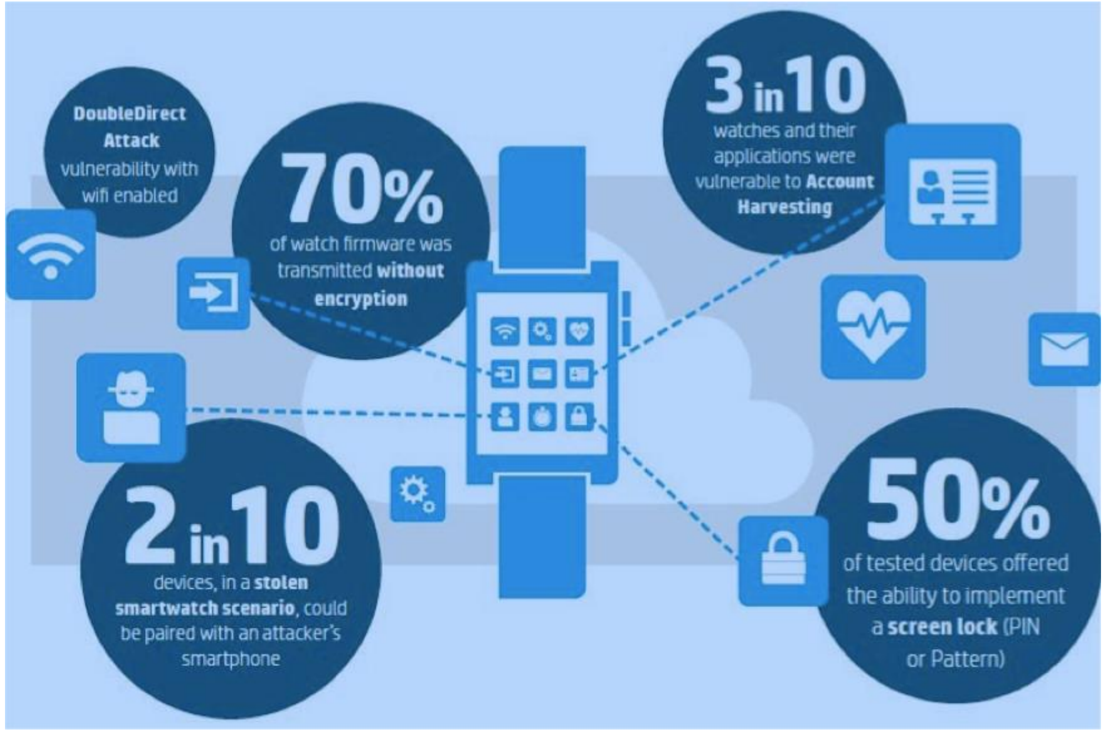
Outline

- PART 1: Motivation
- PART 2: Feasibility
- PART 3: System
- PART 4: Evaluation

1.1: User Authentication



Small form-factor wearables are increasingly **POPULAR** among people

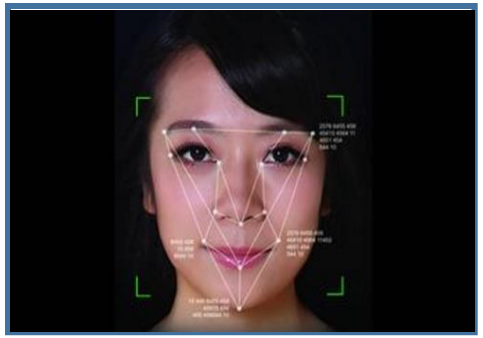


Data privacy issue should be seriously treated for these smart devices

1.2: Existing methods

Fingerprint

Iris recognition

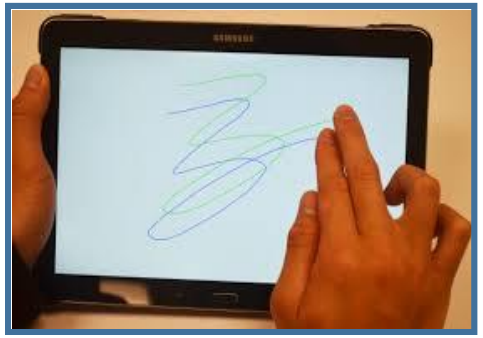


Face recognition

Breath-printing

Voice-printing

Gait recognition



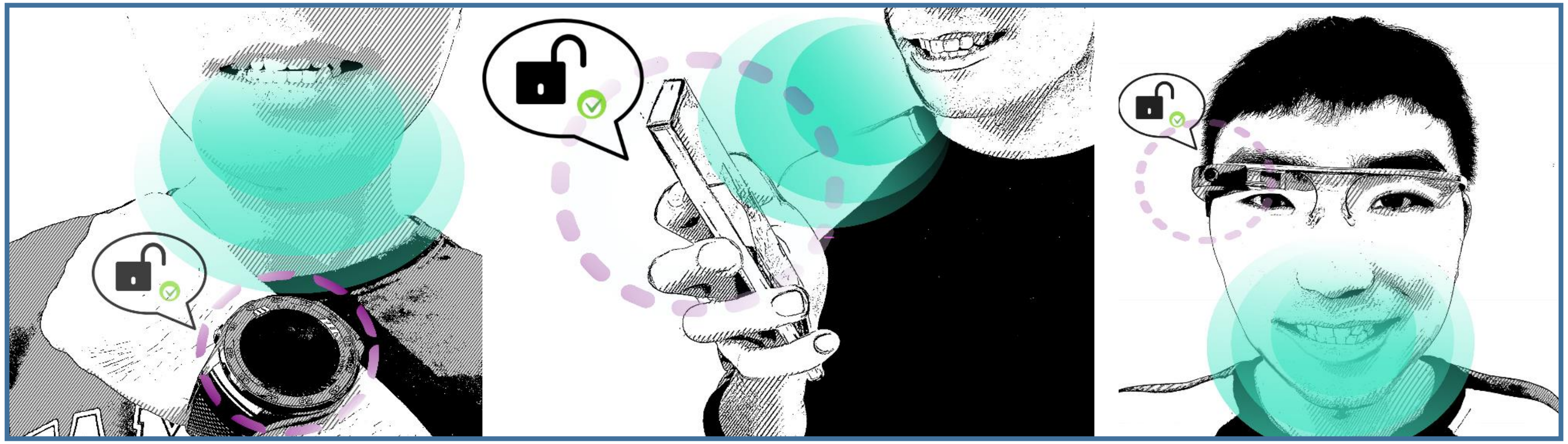
Gesture recognition

Brain wave

1.3: Their limitations

1. **Hardware concern**: sensor size, energy consumption (***Face/Iris/Finger***)
2. **Social acceptance**: feeling embarrassing in public (***Voice***)
3. **Stability** : affected by user's physiologic states (***Breathing/Voice/gait***)
4. **Security:** not robust enough to different kinds of attacks (***Voice***)

1.4: Our proposal

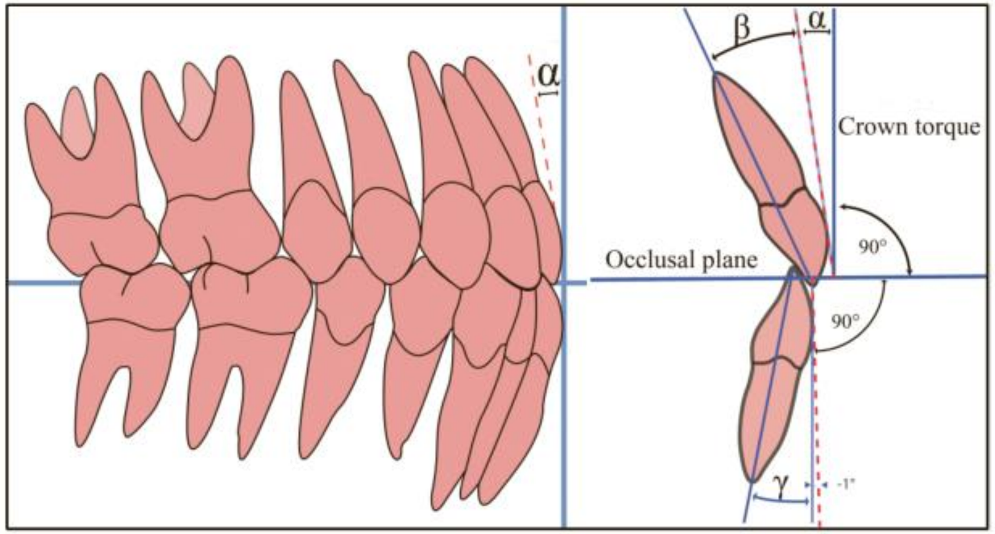


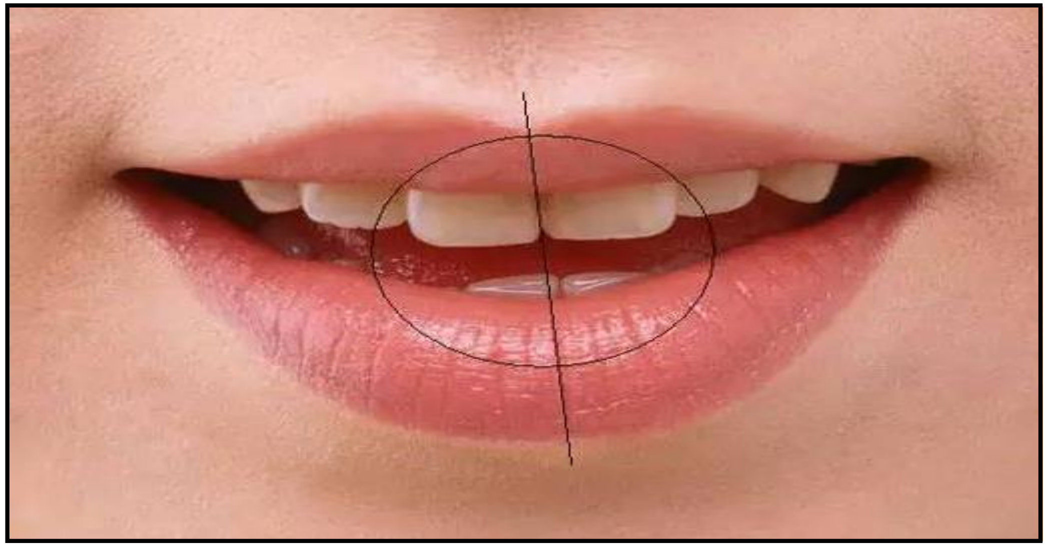
Sounds of tooth click as a biometric for smart devices authentication

Hardware : *pervasive microphone, no additional sensor*
Social acceptance: *more imperceptible and unobtrusive to others*
Stability: *not easily affected by body states*
Security: *robust against replay and observation attacks*

Outline

2.1: Clinic observation





Shape, **Size**, **Orientation** and **Mass** of teeth are different among different people*

**Thomas R Katona and George J Eckert. 2017. The mechanics of dental occlusion and disclusion. Clinical Biomechanics 50 (2017), 84–91.*

2.2: Feasibility study

Hardware

Devices	Class
Samsung Galaxy Tab S2	SM-T815C
Huawei Watch 2	LEO-DLXX
Decibel-meter	AS804
Computer	Hp:498 G3MT
MatLab	2016a

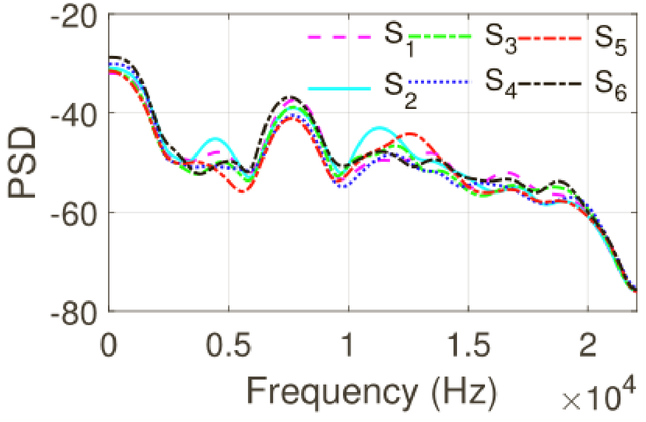
Environment

Meeting room, four noise levels　Lab room, mobile scenario

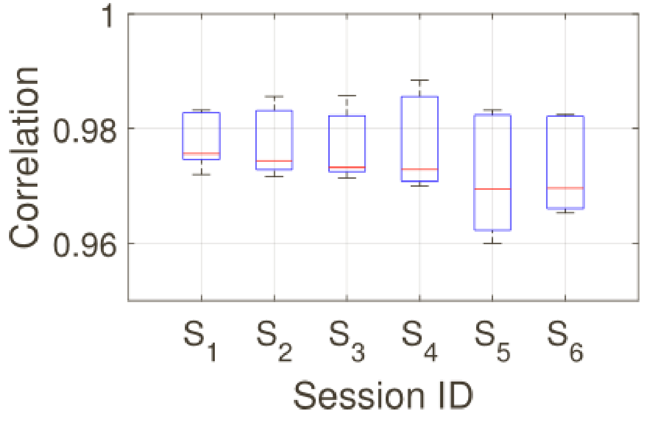
Data collection

- Settings: meeting room (N_1: **30~40** dB, N_2: **40~50** dB, N_3: **50~60** dB, N_4: **60~70** dB), lab room (**40~50** dB**)**
- Sessions: S_1 (**1~2** days, **20** samples**)**, S_2 (**3~4** days, **20** samples), S_3 (**2~3** weeks, **20** samples), S_4 (**1~2** month, **20** samples), S_5 (**3~4** months, **10** samples), S_6 (**5~6** months, **10** samples)
- **D**ata: **100** (number of instances)×**5** (number of settings) ×**50** (number of participants) ×**2** (number of devices)

2.3: Study results



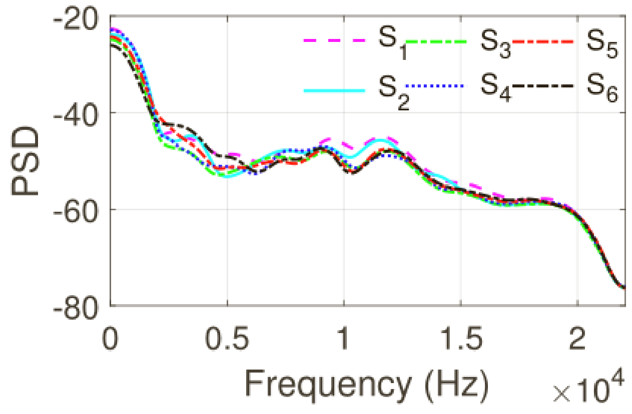
The PSD curve of user X at different time intervals



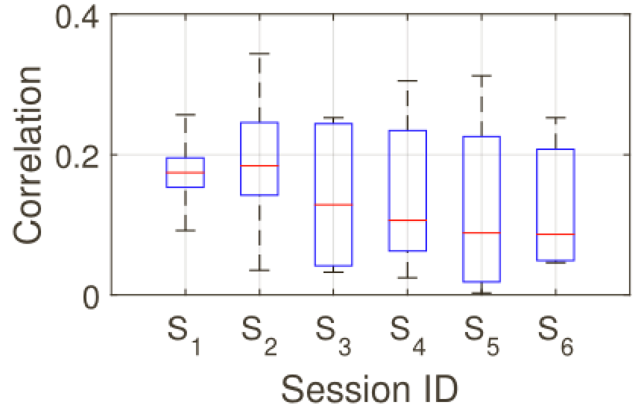
PSD correlation of user X at different time

Notes:

S1: 1~2 day;	S2: 3~4 days
S3: 2~3 weeks;	S4: 1~2months
S5: 3~4 months;	S6: 5~6 months



The PSD curve of user Y at different time intervals



PSD correlation between user X and user Y

Conclusion:

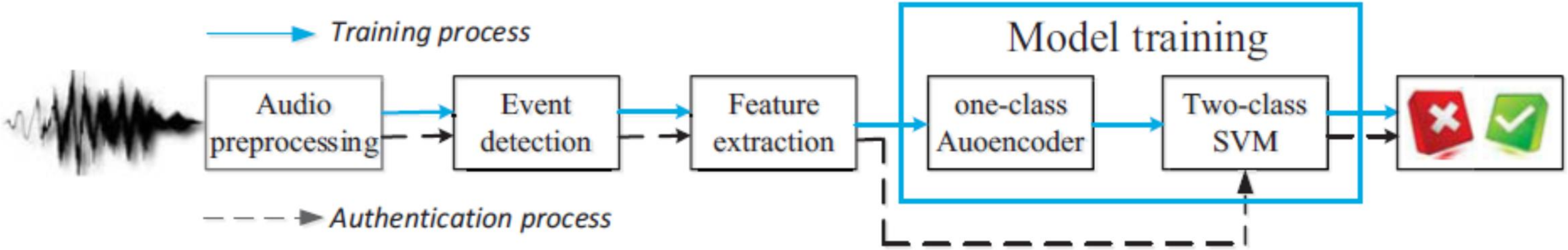
① ***Consistent*** **for the same person**

② ***Different*** **for different persons**

Outline

- PART 1: Motivation
- PART 2: Feasibility
- PART 3: System
- PART 4: Evaluation

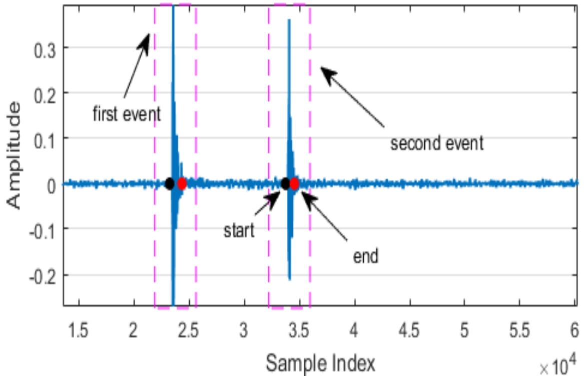
3.1: System architecture



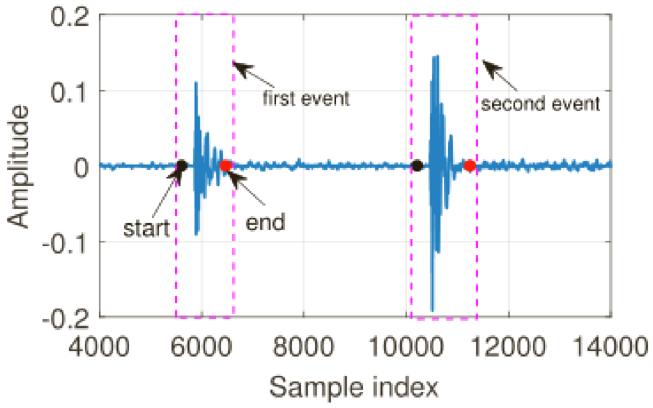
Challenge 1: *how to detect tooth click events adaptively in different environments?*

Challenge 2: *how to design authentication model to accurately authenticate users?*

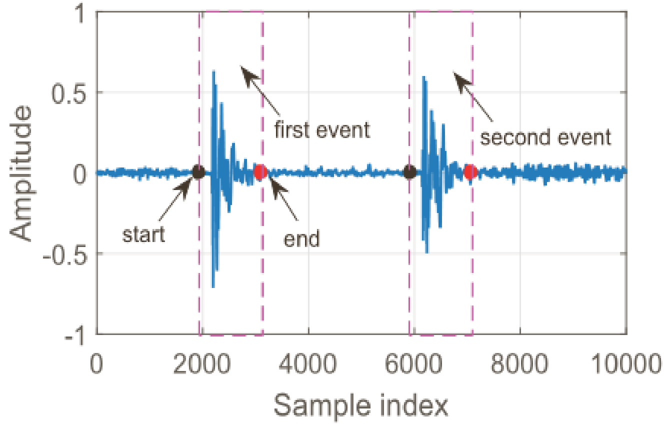
3.2: Event detection



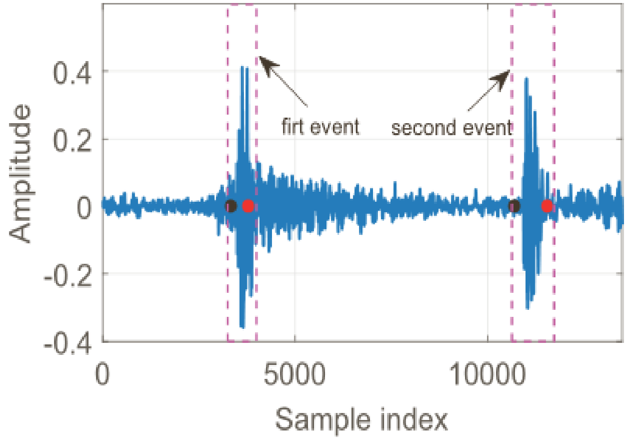
N1: 30-40dB



N2: 40-50dB



N3: 50-60dB



N4: 60-70dB

Improved CFAR:

$$\mu(i) = \frac{1}{W}A(i) + (1 - \frac{1}{W})\mu(i-1)$$

$$\sigma(i) = \frac{1}{W}B(i) + (1 - \frac{1}{W})\sigma(i-1)$$

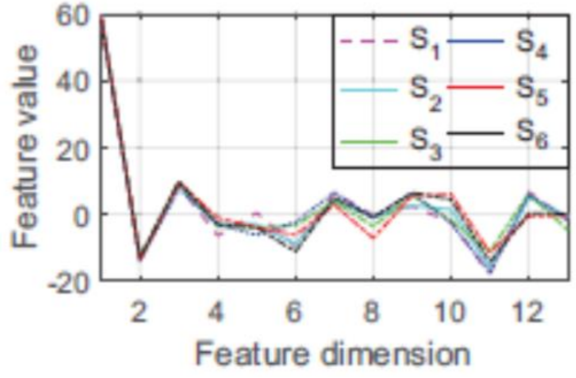
$$A(i) = \frac{1}{W}\sum_{k=i}^{W+i} |S(k)|^2$$

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

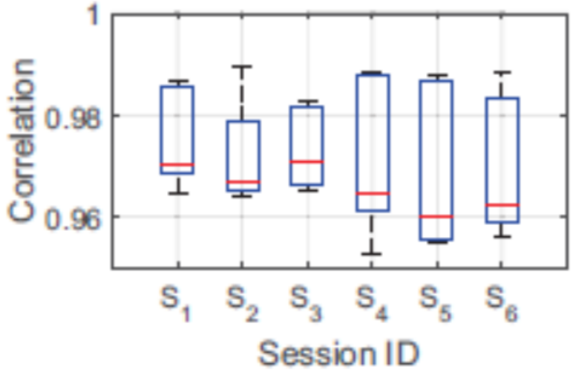
$$|S(i)|^2 > \mu(i) + \gamma_1\sigma(i)$$

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

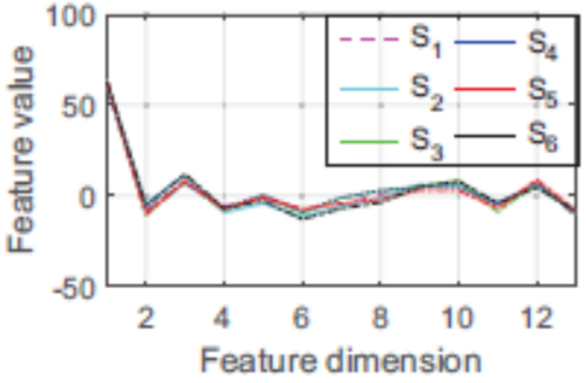
MFCC



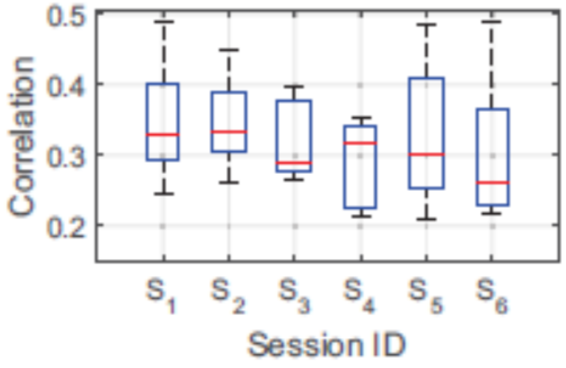
The average feature vector of user X in different sessions



The feature vector correlation coefficients of user X

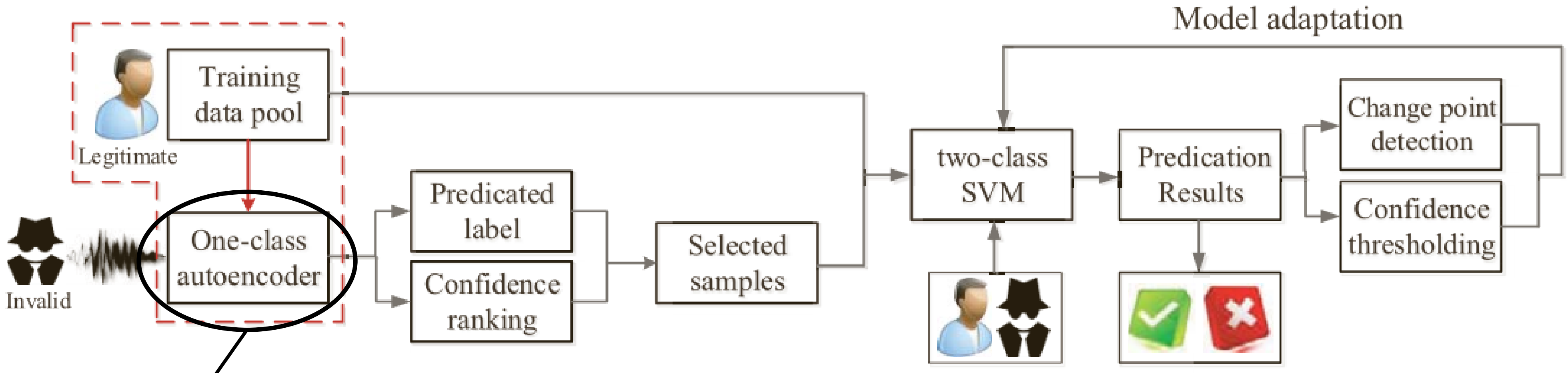


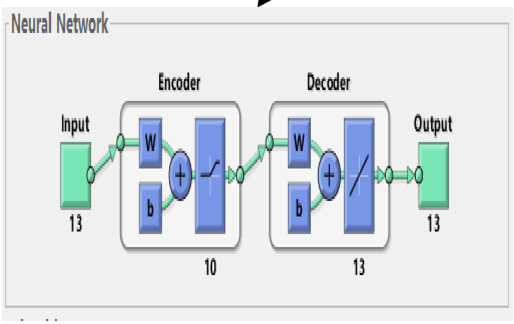
The average feature vector of user Y in different sessions



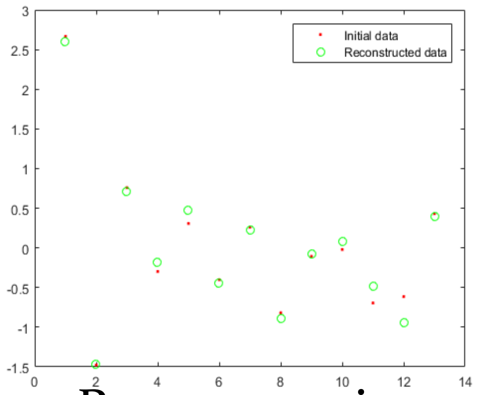
The feature correlation coefficients of user X and user Y

3.3: Model training

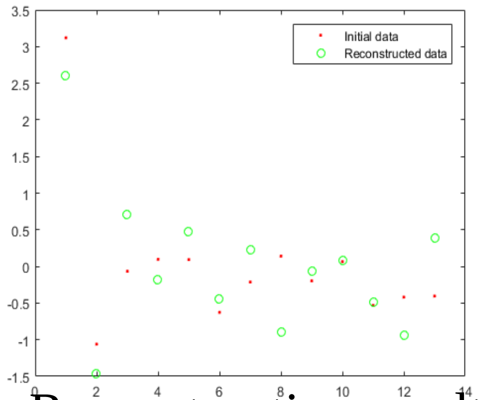




Auto-encoder



Reconstruction result of valid user



Reconstruction result of invalid user

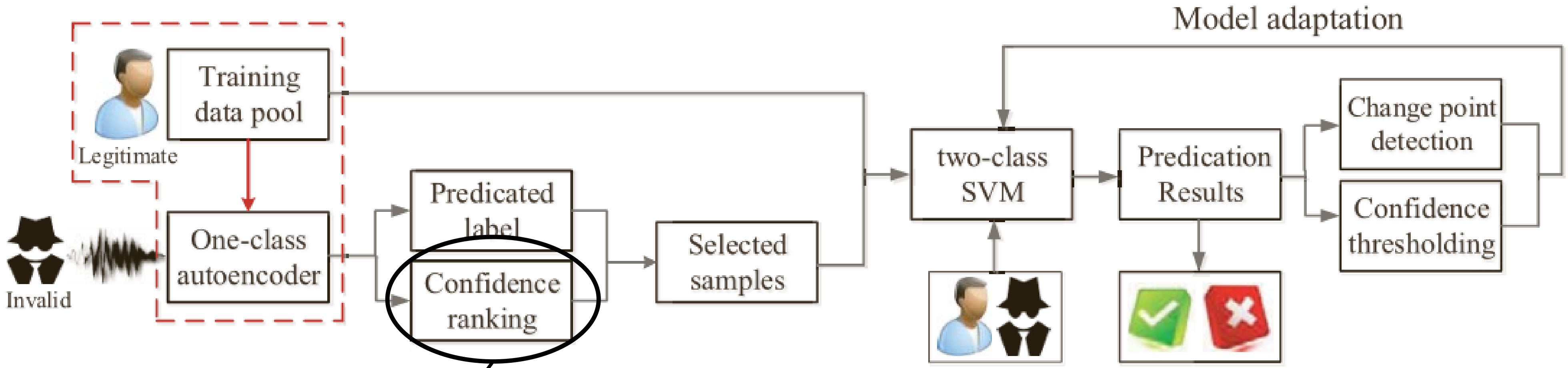
P: max reconstruction error

MSE: reconstruction error

MSE<= **P**, sample of valid user

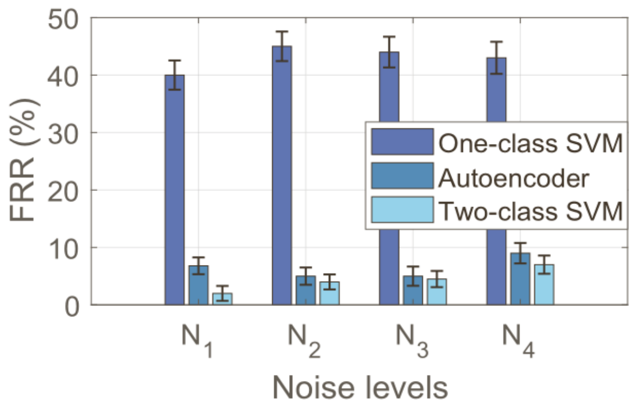
MSE>**P**, sample of invalid user

3.3: Model evolution

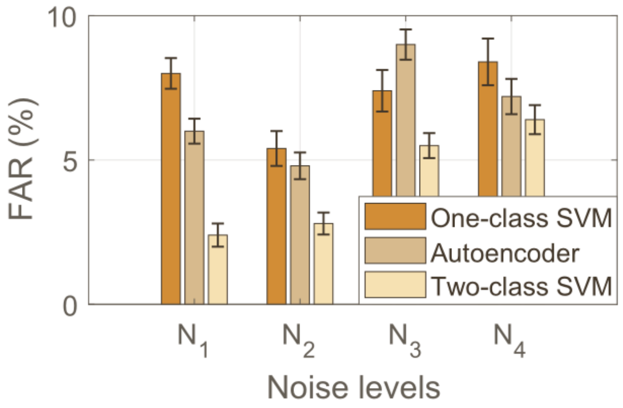


Rank labeled samples by their MSE, and do:

- for positive samples, select samples with small MSE
- For negative samples, select samples with large MSE

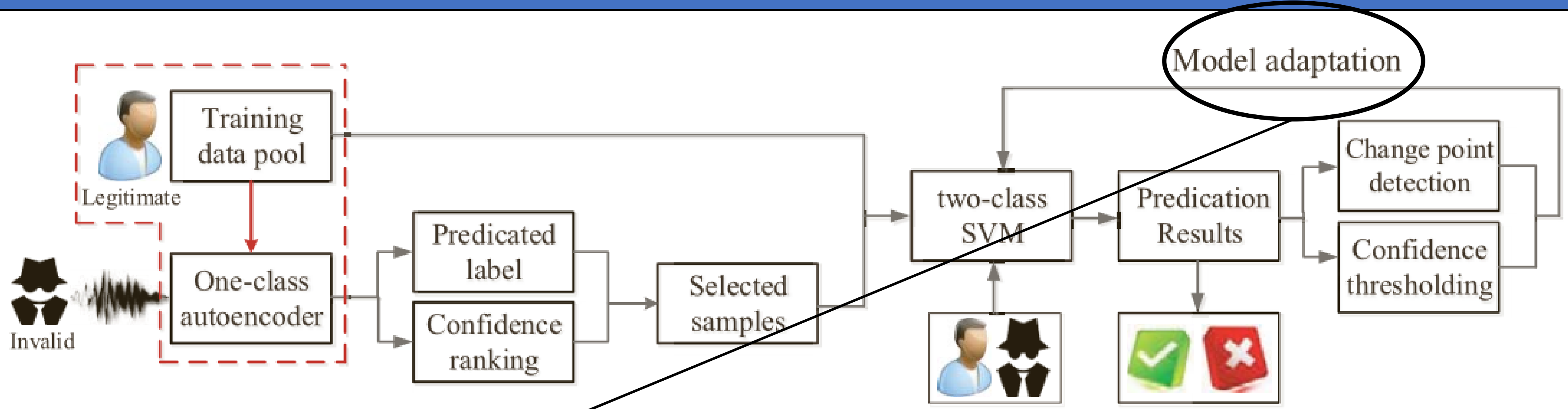


FRR of different methods



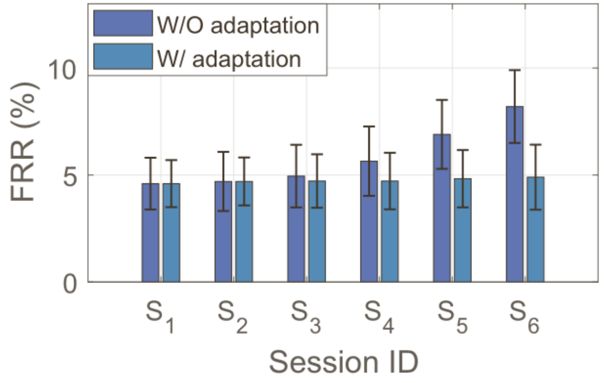
FAR of different methods

3.3: Model adaptation

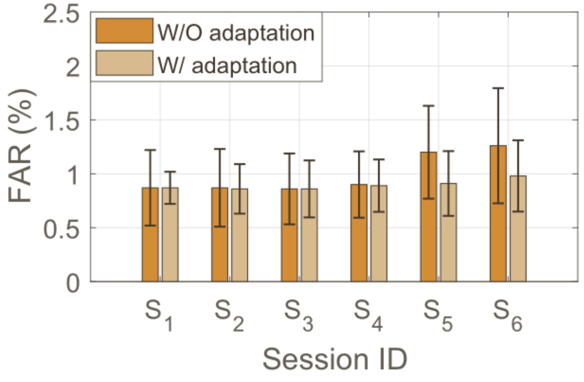


Select samples deviate with previous samples, considering the variation of tooth click in the long term:

Kullback-Leibler (KL) divergence $KL(\Delta t) = \sum_{i=1}^{M} \overline{MFCC_i}^{t} \log \frac{\overline{MFCC_i}^{t}}{\overline{MFCC_i^{t+\Delta t}}}$



FRR



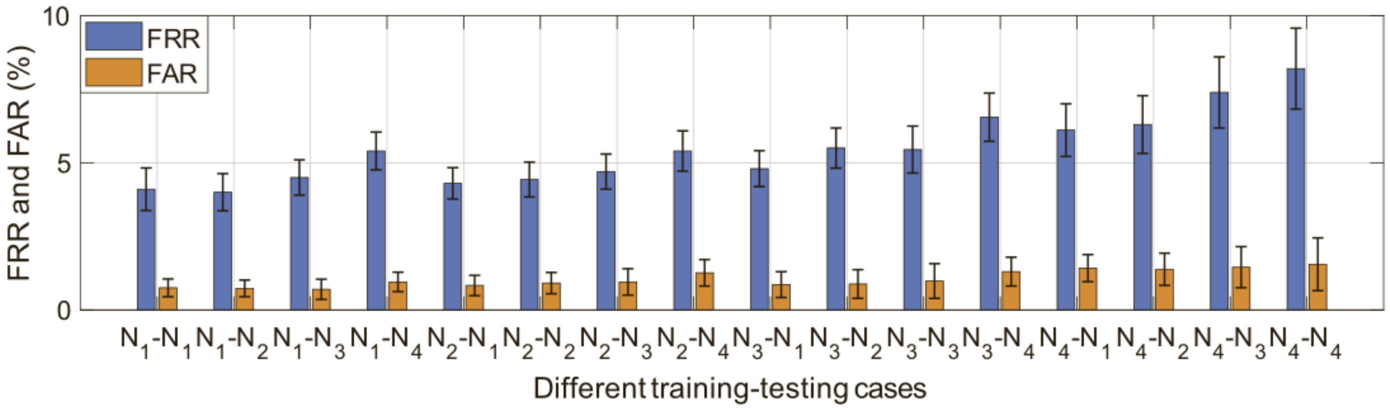
FAR

Outline

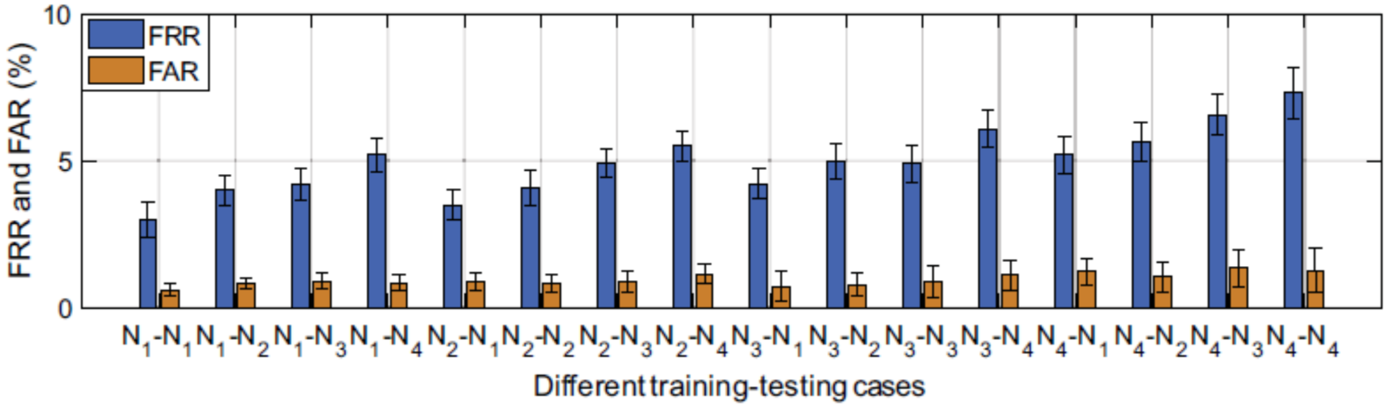
- PART 1: Motivation
- PART 2: Feasibility
- PART 3: System
- PART 4: Evaluation

4.1: Accuracy

Tablet

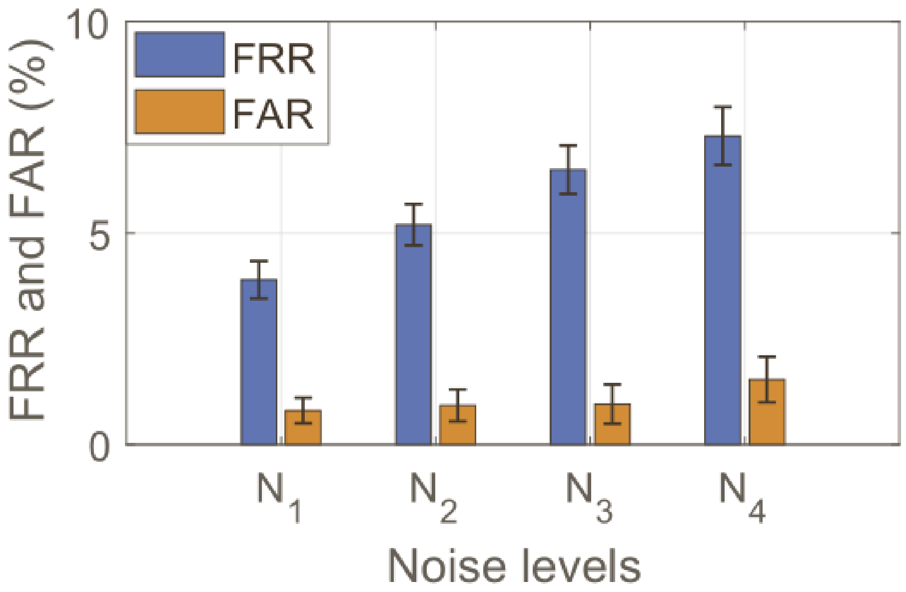


Smartwatch



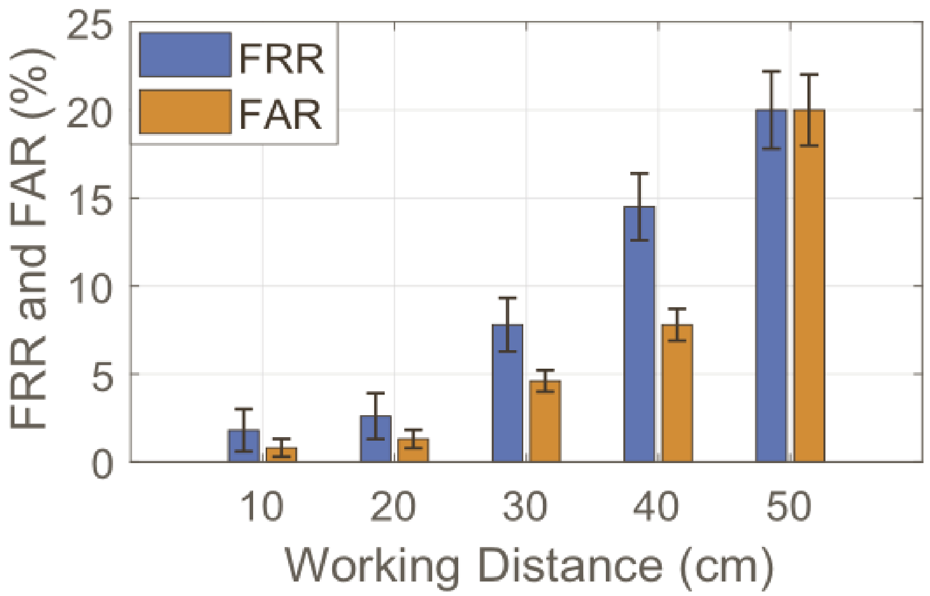
	FAR	FRR
Tablet	1.1%	5.5%
Smartwatch	0.95%	4.5%

4.2: Robustness



Impact of mobility

FAR: **5.7%**, FRR: **1.1%**



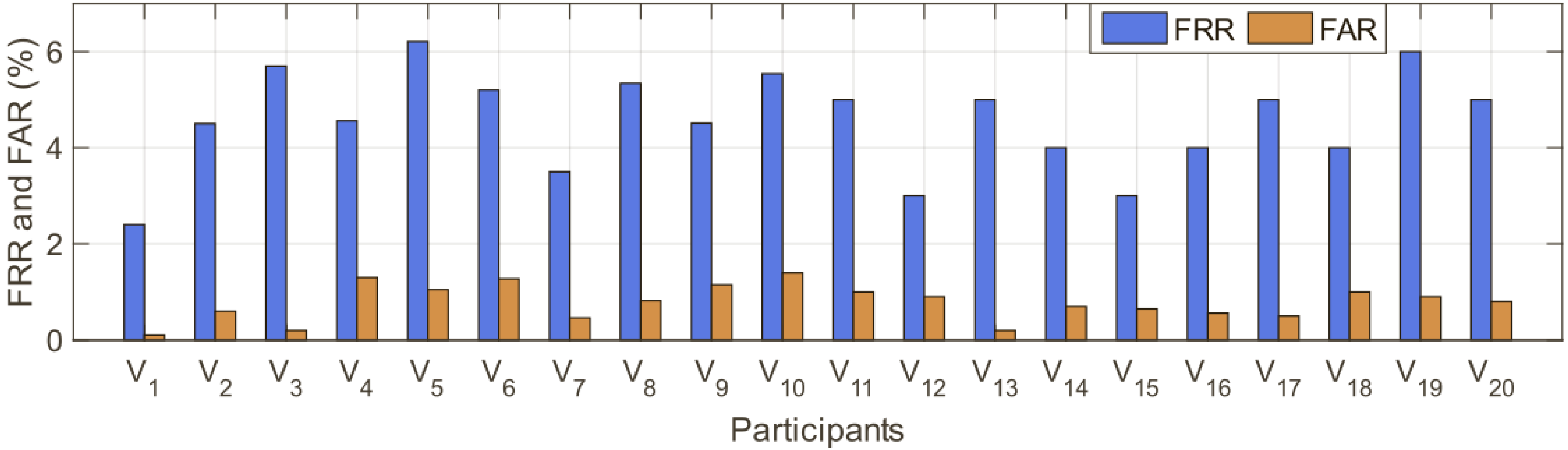
Impact of distance to user's lips

Less than **20 cm**

Mobility nearly **NOT** affect the performance of BiLock

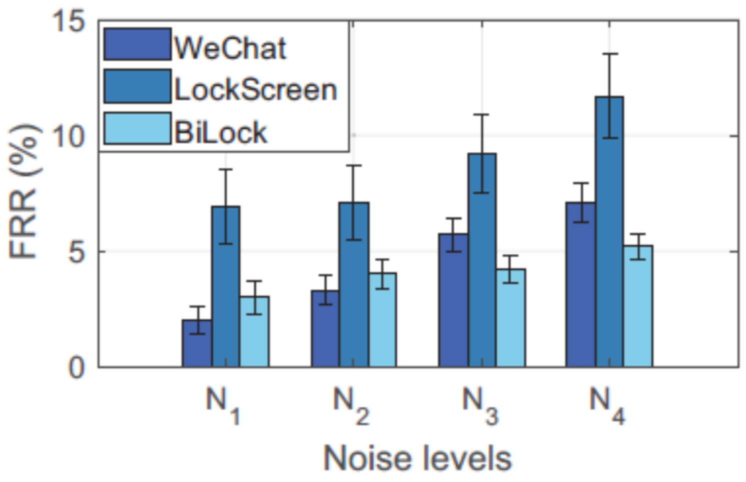
Works well within a distance of less than **20 cm**

4.3: User variance

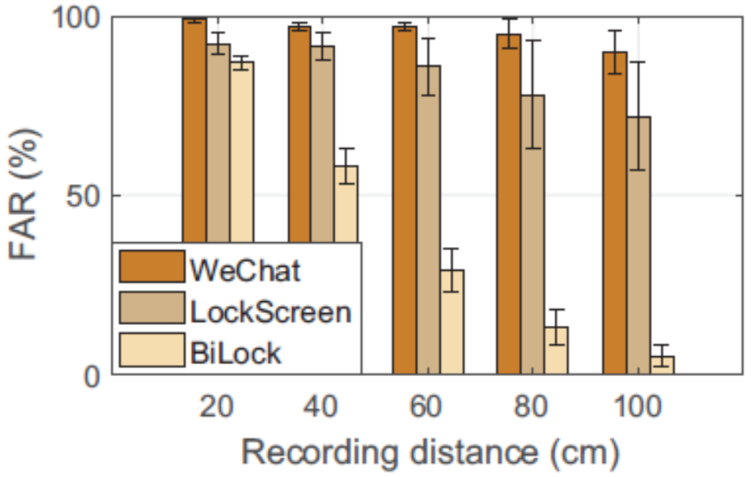


	Max.	Min.
FAR	6.2%	2.4%
FRR	1.4%	1.1%

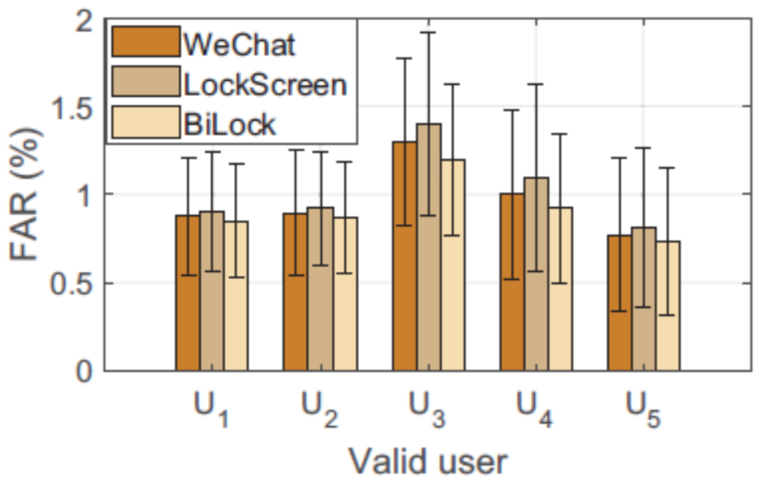
4.4: Comparison



Test WerChat, LockScreen, BiLock under different ***noise levels***



Test WerChat, LockScreen, BiLock under ***replay attacks***



Test WerChat, LockScreen, BiLock under ***observation attacks***

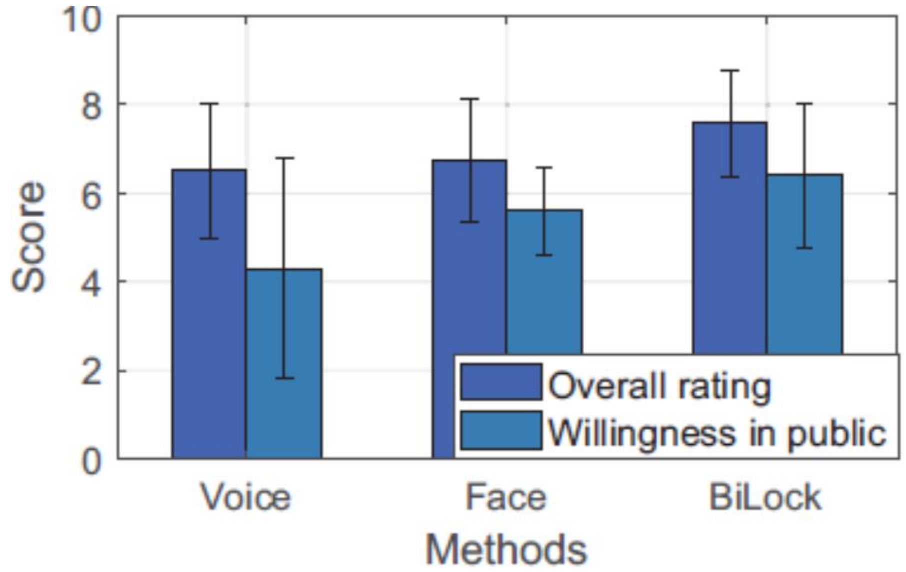
Robustness:
BiLock is ***comparable*** with WeChat, and ***better*** than LockScreen

Replay attack:
BiLock performs ***obviously better*** than other two systems

Observation attack:
BiLock performs ***similarly*** to other two systems

4.5: User experience

100 volunteers, **50** are newly recruited, online questionnaire



$Rating(BiLock) = 7.6$ $Rating(Face) = 6.8$ $Rating(Voice) = 6.5$

$Willing(BiLock) = 6.4$ $Willing(Face) = 5.6$ $Willing(Voice) = 4.3$

Nonparametric Wilcoxon signed-rank test for rating:

$Z(BiLock, Voice) = -2.27$ $\mathrm{p}(BiLock, Voice) = 0.012$

$Z(BiLock, Face) = -1.79$ $\mathrm{p}(BiLock, Face) = 0.037$

- *"It is rather embarrassing to speak out words in public when using voiceprinting method. In contrast, BiLock is more imperceptible and easy to use. But I prefer to use BiLock without placing the device so near to my mouth if possible."*
- *"I use voice-prints frequently but BiLock is also cool. I think BiLock may be more robust when I caught a cold. Sometimes my phone does not recognize my voice when I got sick."*

Conclusion

We propose a novel biometric authentication scheme with good ubiquity, high robustness and security based on human tooth clicks

We design methods to extract tooth click events adaptively in different environments, and effective authentication model with self-adaptation

The experimental results show that in the normal noise environment of 50~60 dB, th authentication recognition model achieves FRR less than 5.0%, FAR less than 0.95%.

THANK YOU

YONGPAN ZOU
Shenzhen University
yongpan@szu.edu.cn
https://yongpanzou.github.io/