Mobile Phones Know Your Keystrokes through the Sounds from Finger’s Tapping on the Screen

Zhen Xiao¹, Tao Chen¹, Yang Liu¹, Zhenjiang Li¹,²
¹Department of Computer Science, City University of Hong Kong
²City University of Hong Kong Shenzhen Research Institute
Tapping sound on the mobile phone

Privacy:
- Message
- Password
- Bank account
Infer keystrokes

Privacy leakage
Distinct tapping sound

- Heterogeneous structure
- Distinct sound
Hacking acoustic signal
Existing methods

- Microphones and IMU sensors
- Fake keyboard

Stealthiness

[WiSec,12] [WiSec,14]
Challenge-I: keystroke recognition

- **Difficult** task with **weak** signal
Challenge-II: unlabeled data

How to classify?

Unlabeled data
System overview

Victim’s phone

Signal pre-processing and segmentation

TDoA for pregrouping

Auto-encoder

Clustering

Unsupervised keystroke recognition

Keystrokes
Keystroke recognition: weak signal

- **Sampling**
- **Raw signal**
- **Wiener Filter**
- **Filtered signal**

- **Start point**
- **End point**

- **Noisy segmentation**
  - 40 ms

- **Clear segmentation**
  - 40 ms

- **High SNR**
- **Clear boundary**
Keystroke recognition: pre-grouping

\[ \Delta t = \frac{(d_t - d_b)}{v} \]

Connect points with the same TDoA

192kHz sampling rate
Keystroke recognition: pre-grouping

\[ \text{Error} = (T_t - T_b) - (\hat{T}_t - \hat{T}_b) \]
Unsupervised classification: clustering

MFCC feature

Low accuracy?
Need better features
Auto-encoder based clustering

\[ L_{\text{loss}} = L_{\text{rec}} + \alpha L_{\text{inner}} - \beta L_{\text{inter}} \]
Experiment setup

- Participants: 6 volunteers as victim users

- Dataset:
  - 200 keystrokes of each key from the adversary for training
  - 4680 keystrokes from the victim users for inference

- Training: Intel i7-8700K CPU and Nvidia GTX 2080Ti GPU

- Testing: Samsung GalaxyS7, Nexus 5X and Huawei P30 Pro
Evaluation

• Overall performance

Average accuracy (top-3): 92.2%
Evaluation

- System Component

- Accuracy: (top-3)
  - TapLeak: 92.9%
  - No pre-grouping: 74.1%
  - No auto-encoder: 63.5%

Impact of system component
Evaluation

- Different users

- Different phones

- Different noise levels

- Robust for different situations
Conclusion

• Demonstrate a possible privacy leakage through dual microphones on the mobile phone

• Propose effective unsupervised model to infer keystrokes with weak acoustic signal

• Implement the prototype of the system and evaluate on different mobile phones