KeyListener: Inferring Keystrokes on QWERTY Keyboard of Touch Screen through Acoustic Signals

Li Lu*, Jiadi Yu*, Yingying Chen†, Yanmin Zhu*, Xiangyu Xu*,
Guangtao Xue*, Minglu Li*

*Dept. Computer Science and Engineering, Shanghai Jiao Tong University
†WINLAB, Rutgers University

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

RUTGERS

Common but Valuable Typing Behavior

- Typing Behavior
  - Common:
    - Widely-used electronics (e.g., PC, smartphone) require keystroke typing as a input method
  - Valuable:
    - About 43% users in the USA adopt mobile banking and typing password for their daily financial activities in 2015 (Federal Reserve System)
    - Around 1,500 million users chat online monthly through instant messaging APPs on smartphones

Vulnerable Typing on a Physical Keyboard

Acoustic

Camera

Wi-Fi

Electromagnetic Radiation

Motion Sensors

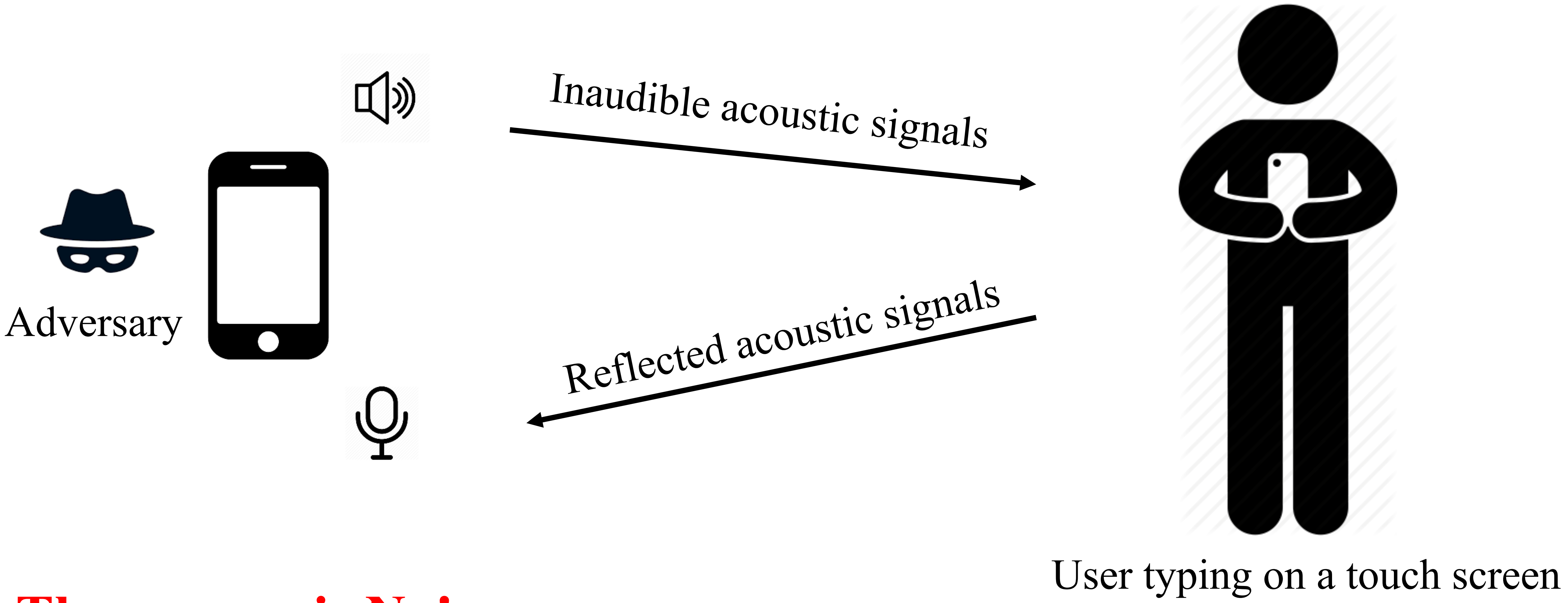
Typing is exposed to various attacks!

What about Typing on a Touch Screen?

- Virtual keyboard
- Tiny finger movements
- No obvious click sound
- Cautious users
- ……

Is that secure enough?

Side-channel Attack using Commercial Devices



The answer is No!

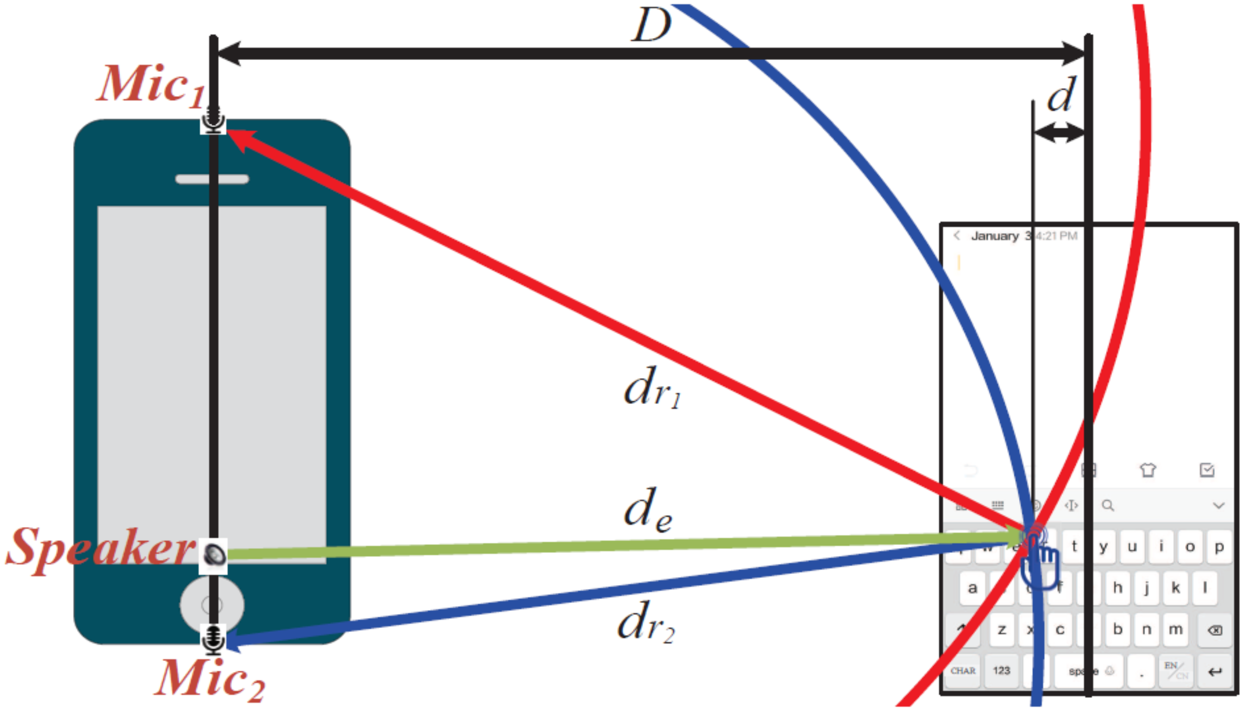
Outline

- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Basic Idea

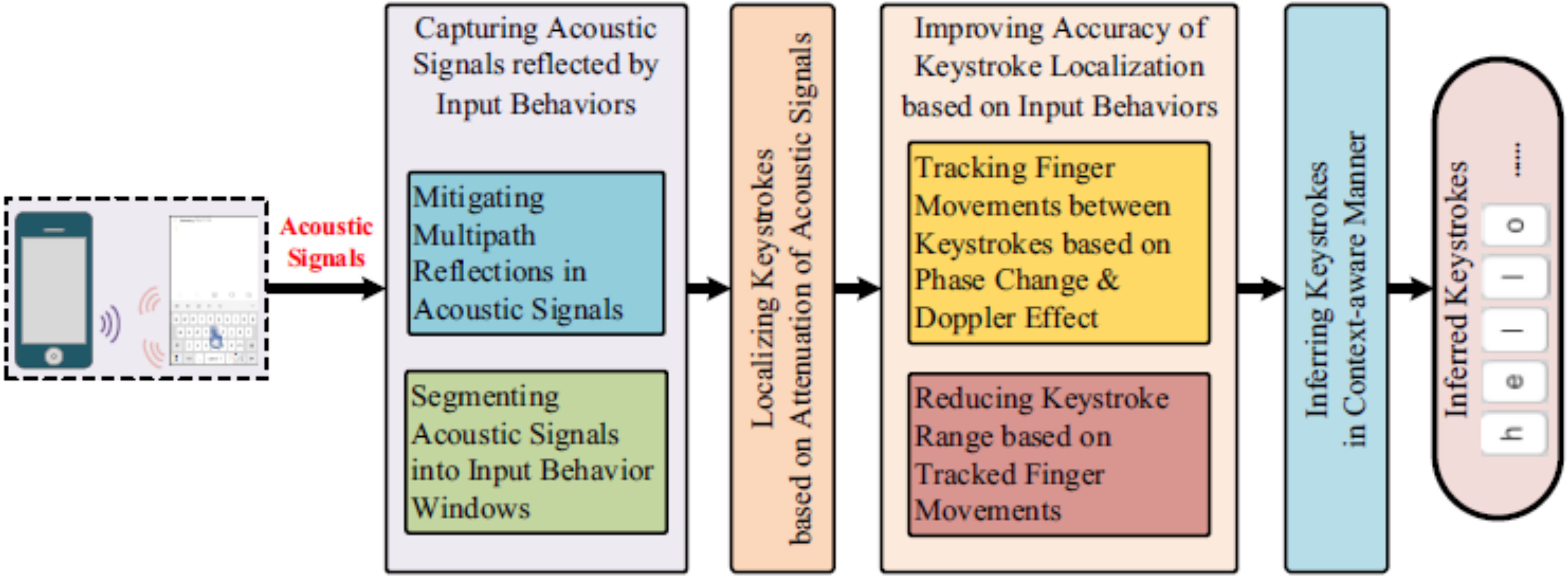
Localize keystroke position in typing behavior leveraging the **attenuation of acoustic signals**:

$$I_r = I_e \frac{k}{d} e^{\alpha d}$$



- Low-cost audio infrastructures in commercial smartphones
- Easily accessed by a curious or malicious adversary
- Hard to be aware by a touch-screen typing user

System Architecture



Outline

- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

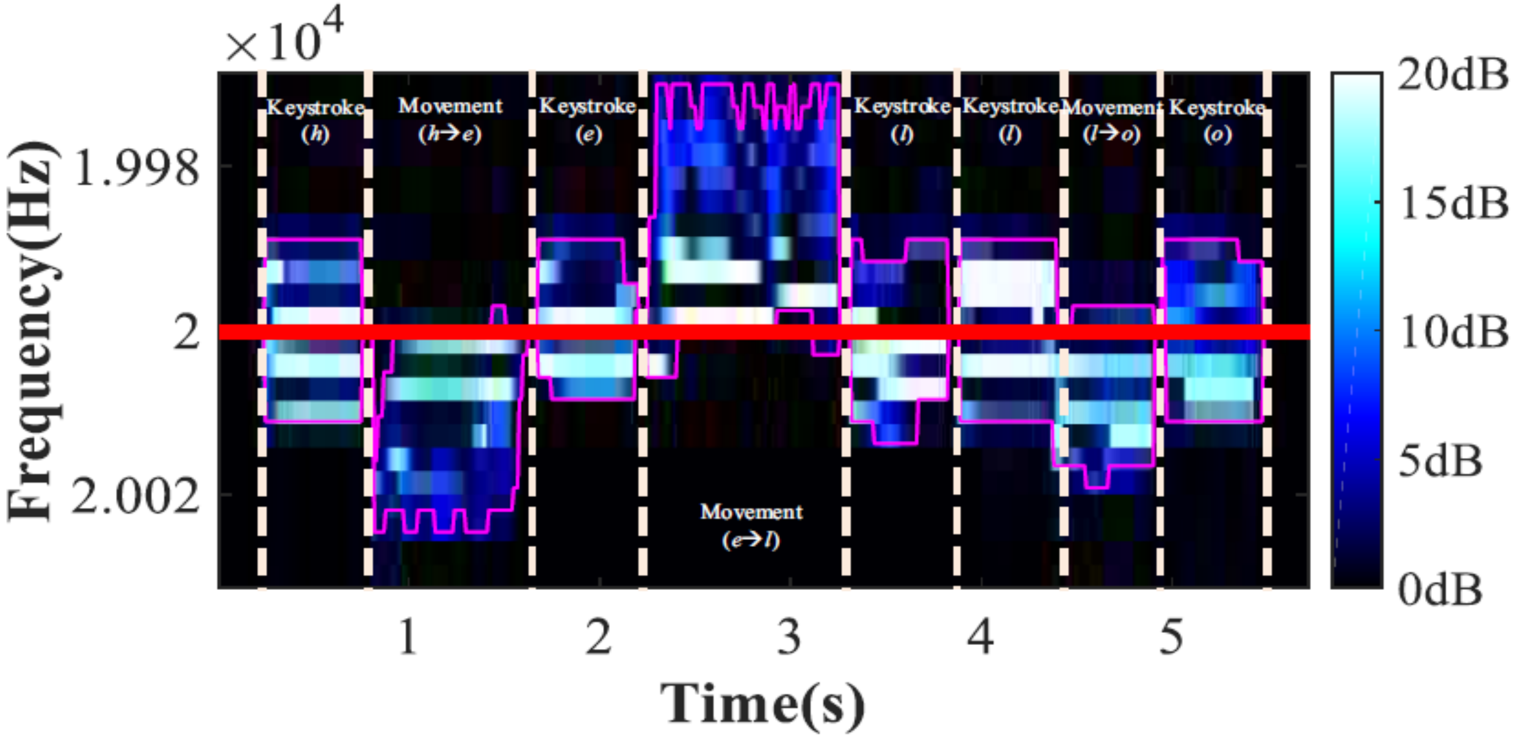
Capturing Signal Reflected by Input Behaviors

- **Mitigating Multipath Reflections in Acoustic Signals**
  - Eliminate LOS signal by computing difference between successive time slots: $g(t) = s(t) - s(t-1)$
  - Mitigate multipath reflections from other dynamic objects by using FFT power: $I_r(t) = \sum_{f=f_0-\Delta f}^{f_0+\Delta f} g(t)$.

- **Segmenting Acoustic Signals into Input Behavior Windows**
  - Separate keystrokes and finger movements using Doppler shift
  - Segment each keystroke and finger movement window

Capturing Input Behaviors

- Example
  - Separate keystroke behavior with finger movements behavior



Spectrogram of received signal when a victim inputs ‘hello’

Outline

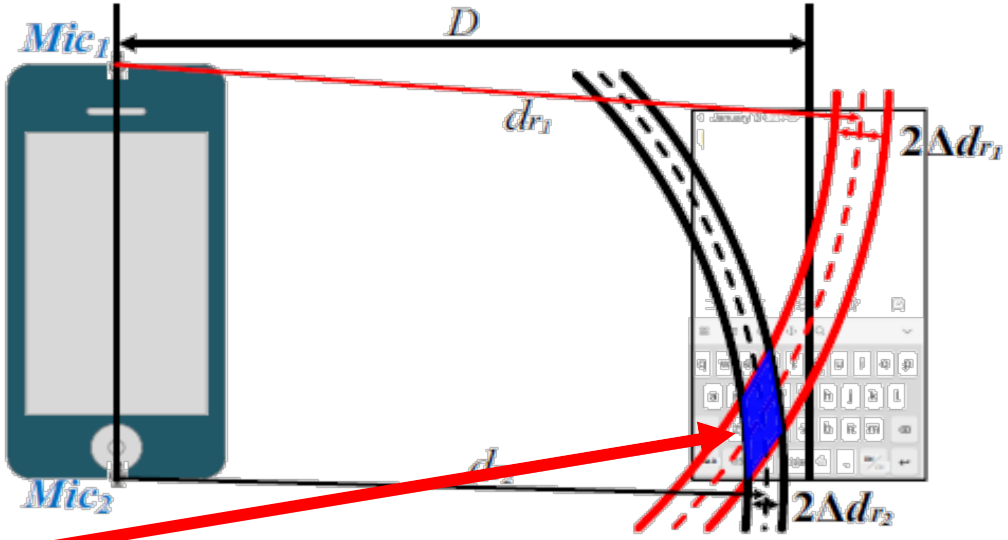
- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Localizing Keystrokes

- **Localizing keystrokes based on attenuation of acoustic signals**
  - Ambient noises matters:

$$I_r \pm I_n = I_e \frac{k}{d \mp \Delta d} e^{\alpha(d \mp \Delta d)}$$

  - Induce significant errors in localizing keystroke positions
  - Localizing to an area, i.e., keystroke range



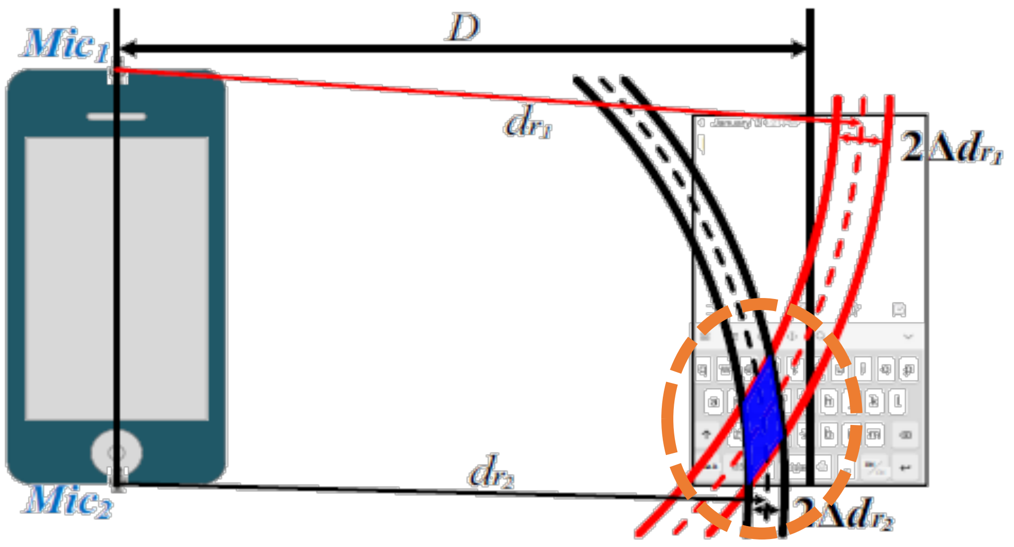
Locate an area as keystroke range

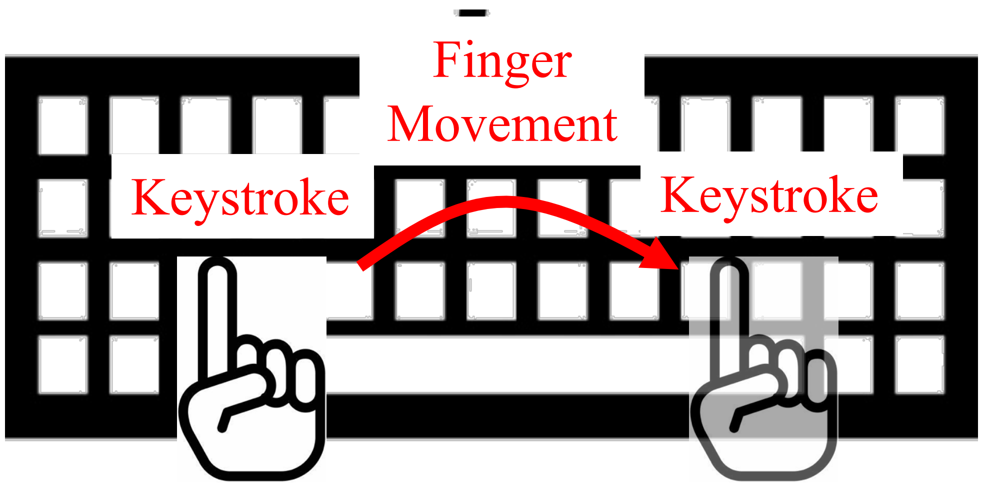
Outline

- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Improving Localization Accuracy

- **Localizing Keystroke Position**
  - Still exist significant errors in localizing keystroke positions
  - Need to be improved!
- **Intuition for Improving Localization Accuracy**
  - Typing Behavior = Keystroke + Finger Movement
  - Utilize finger movement to improve localization accuracy





Improving Localization Accuracy

- **Tracking Finger Movements based on Phase Change and Doppler Effect**
  - Tracking finger movement distance between keystrokes using Phase Change

1. Emitted acoustic signal: $s_e(t) = A\cos(2\pi f_0 t)$
2. Received acoustic signal: $s_r(t) = A'\cos(2\pi f_0 t - \frac{2\pi f_0 d}{c})$
3. Multiply: $s_r(t) \times s_e(t) = \frac{1}{2}AA'(\cos\left(-\frac{2\pi f_0 d}{c}\right) + \cos(4\pi f_0 t - \frac{2\pi f_0 d}{c}))$
4. Low-pass filtering: $\frac{1}{2}AA'\cos\left(-\frac{2\pi f_0 d}{c}\right)$
5. Distance calculation: $d = -\frac{\phi_t - \phi_0}{2\pi} \times \frac{c}{f_0}$

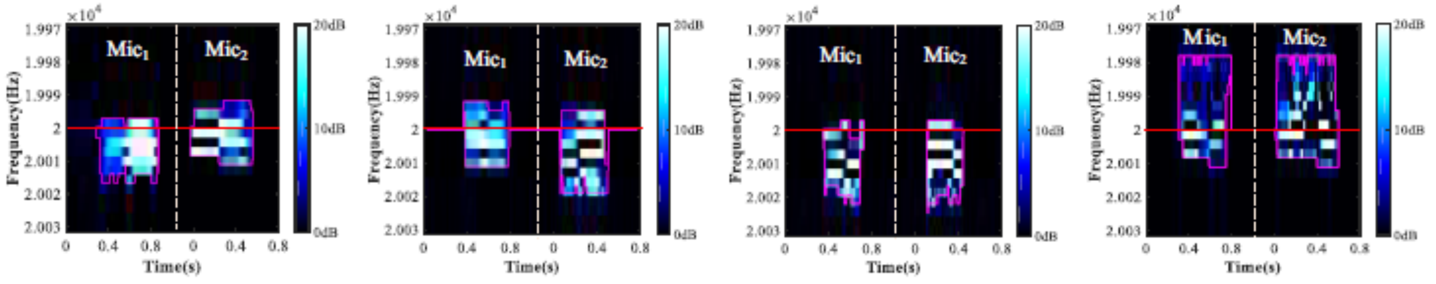
ϕ_t is the phase value in time t

c is the speed of acoustic signal

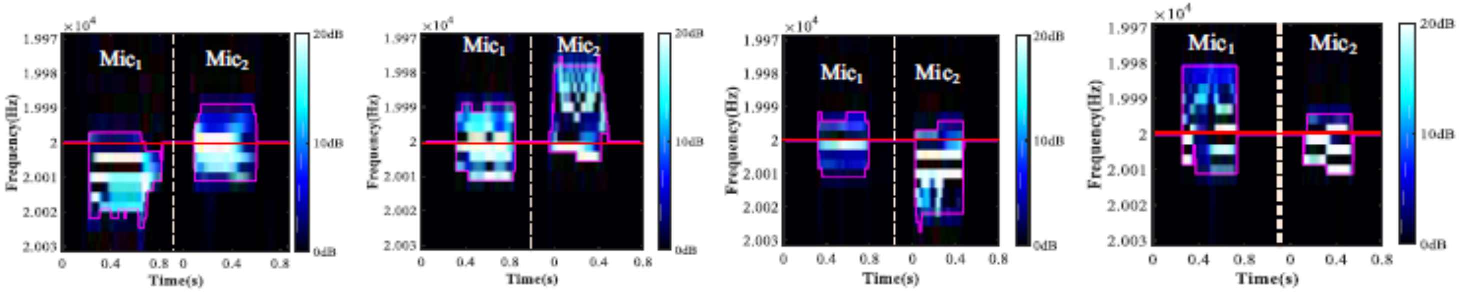
f_0 is the frequency of pilot tone (20kHz in our system)

Improving Localization Accuracy

- **Tracking Finger Movements based on Phase Change and Doppler Effect**
  - Tracking finger movement direction between keystrokes using Doppler Effect



(a) Up. (b) Down. (c) Left. (d) Right.

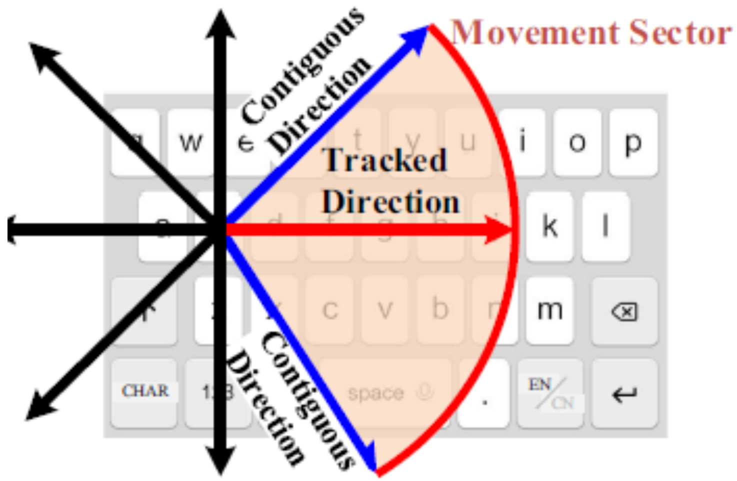


(e) Top-Left. (f) Top-Right. (g) Bottom-Left. (h) Bottom-Right.

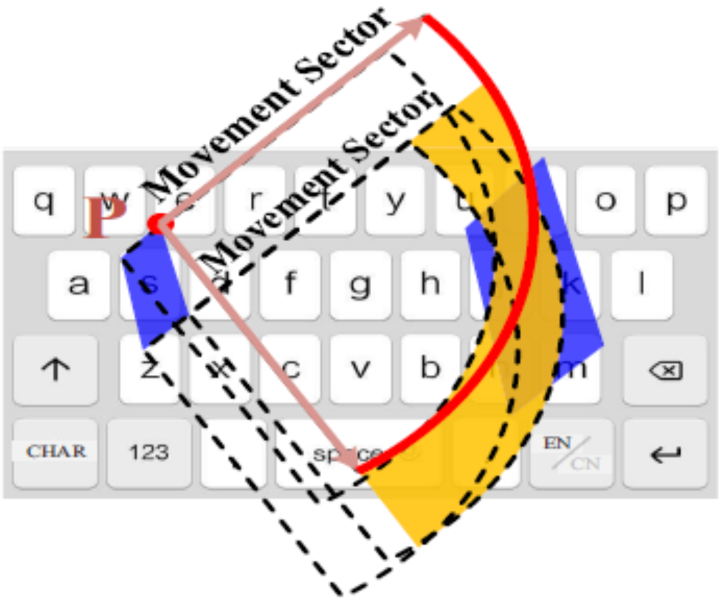
Doppler patterns of eight different directions

Improving Localization Accuracy

- **Reducing Keystroke Range based on Tracked Finger Movements**
  - Constructing movement sector based on localized keystroke and tracked finger movements
  - Reducing keystroke range with the movement sector



(a) Movement sector.



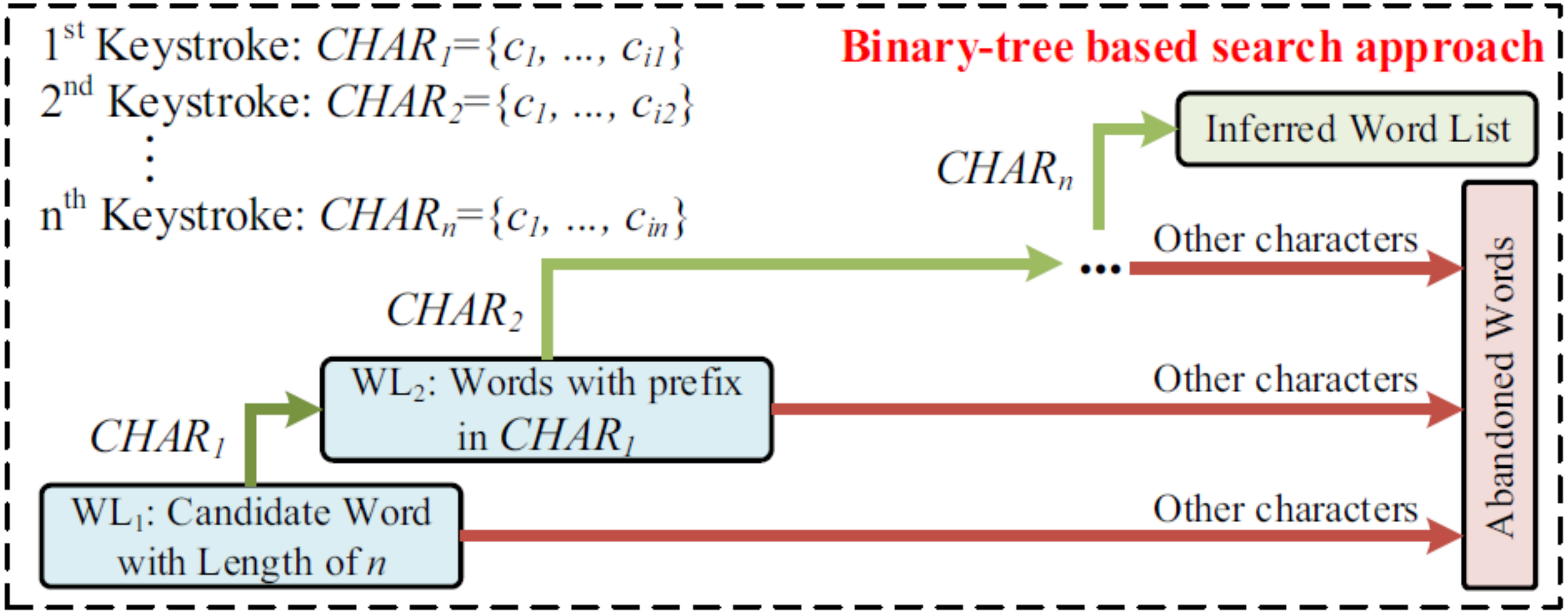
(b) Keystroke range reduction.

Outline

- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Inferring Keystrokes in Context-aware Manner

- **Binary tree-based search approach**
  - Still cannot localize precisely to a key on the keyboard
  - Exist multiple character candidates for one keystroke localization
  - Utilize context information during input to infer the input



Outline

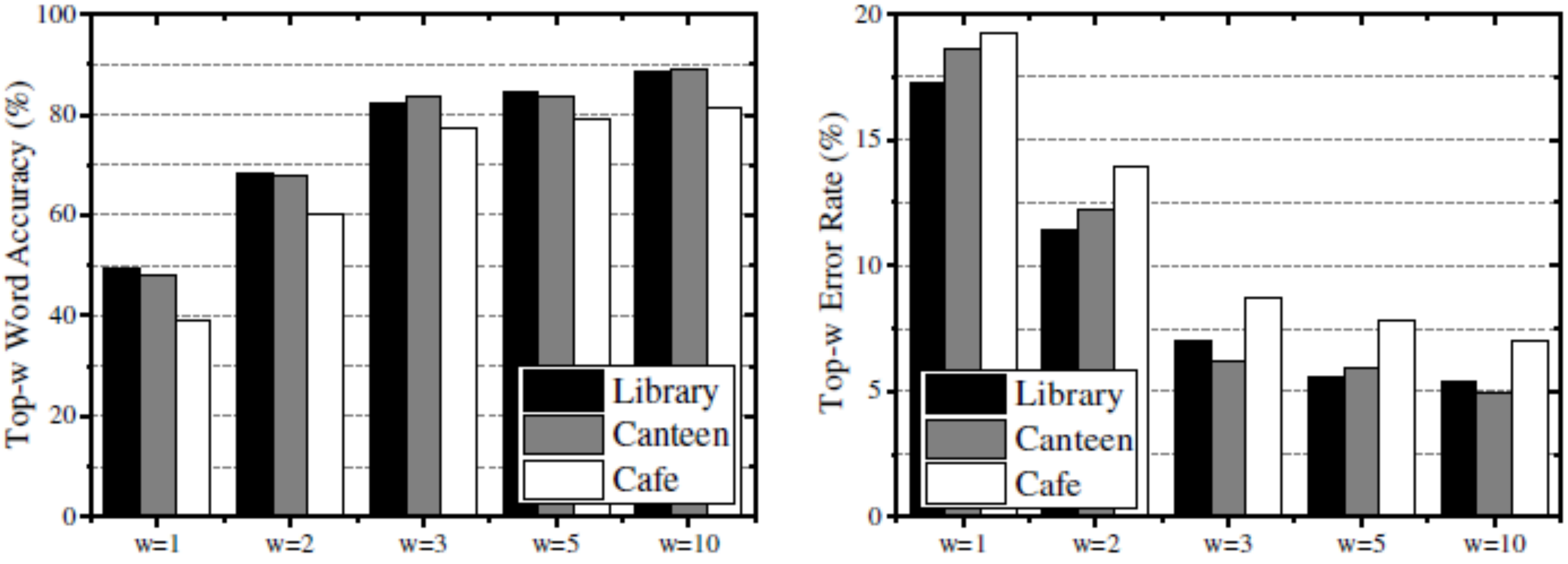
- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Experiment Setup

- Adversary: implementation of **KeyListener** on a Galaxy S4 with Android 5.1.1
- Victim user:
  - 24 volunteers, 12 males and 12 females with ages in [18, 45]
  - Four types of user smartphones: 4.7-inches iPhone 7, 5.2-inches Huawei P7, 5.5-inches iPhone 7 Plus and 7.0-inches Huawei Honor X2
- Environments
  - Scenario: 1. sitting in a library 2. sitting in a canteen 3. queuing in a café
  - Placement of adversary's smartphone relative to victims: left, right and opposite
  - Distance between smartphones of the adversary and victim: 45~60 cm

Overall Performance

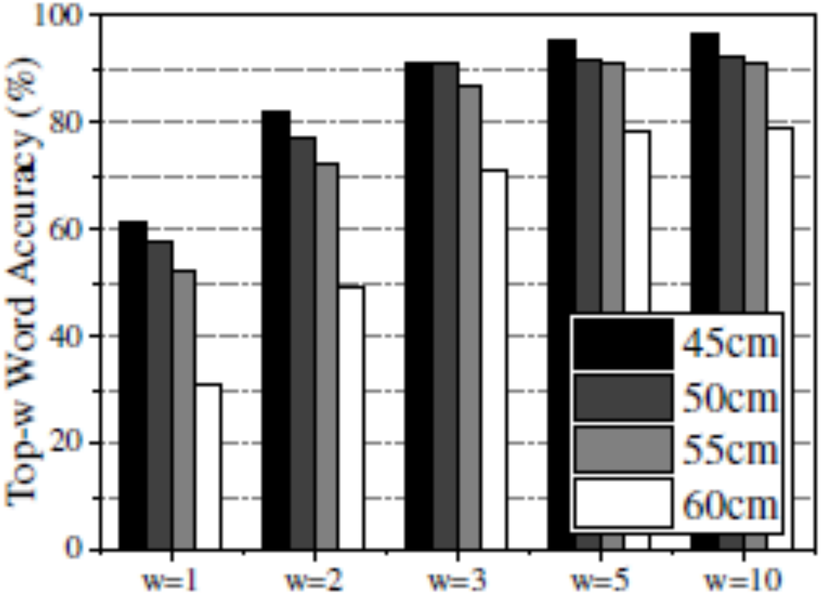
Top-10 word accuracies in the library and canteen can approach 90%, in the cafe is 81.3%; top-w error rates are satisfactory.



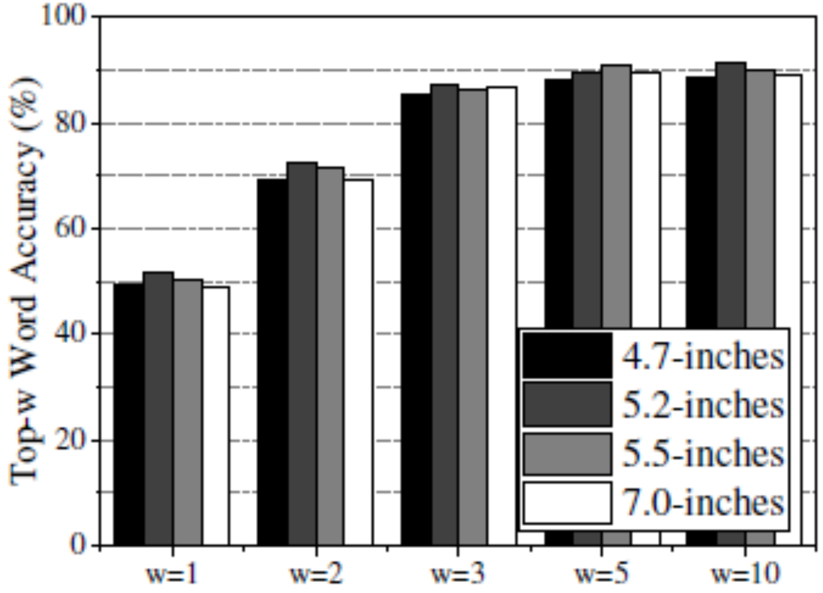
(a) Top-w word accuracy. (b) Top-w error rate.

Impacts of Different Factors

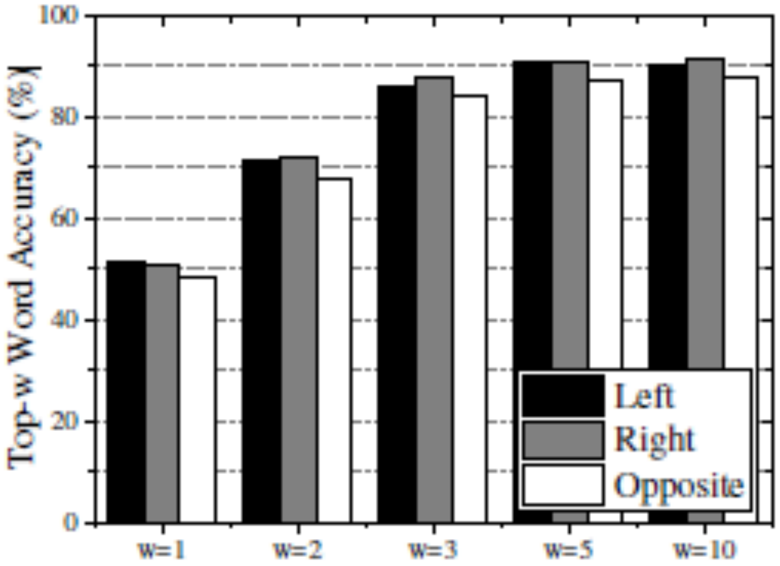
Distance



Screen size



Relative position

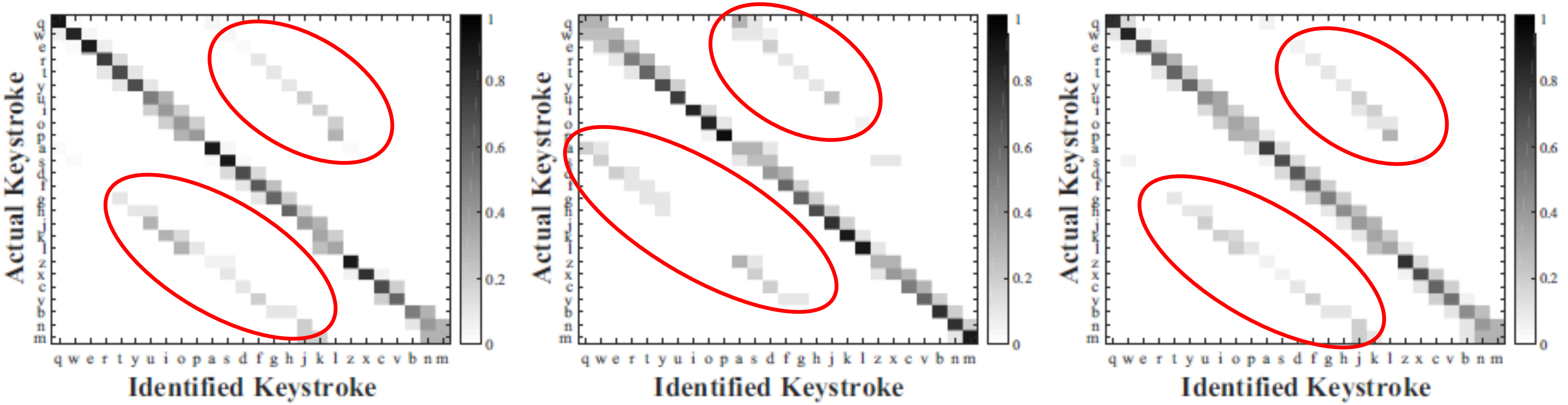


High accuracy within 60cm

Accurate under different screen size

Opposite is a little lower due to obstacle

Single Keystroke Identification



(a) Left. (b) Right. (c) Opposite.

Relative position matters!

Outline

- System Design
  - Capturing Input Behaviors
  - Localizing Keystrokes
  - Improving Localization Accuracy
  - Inferring Keystrokes in Context-aware Manner
- Evaluation
- Conclusion

Conclusion

- Revealing a side-channel attack based on acoustic signals by commercial smartphone
- Localizing keystrokes based on attenuation of acoustic signals
- Improving the keystroke localization accuracy through tracking finger movements between two successive keystrokes
- Extensive experiments demonstrate that KeyListener could achieve sufficient accuracy for keystroke snooping on QWERTY keyboard of touch screen

Thank you!

Q & A

上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

RUTGERS