When CSI Meets Public WiFi: Inferring Your Mobile Phone Password via WiFi Signals

Mengyuan Li[1], Yan Meng[1], Junyi Liu[1], Haojin Zhu[1], Xiaohui Liang[2],
Yao Liu[3] and Na Ruan[1]

[1]Shanghai Jiao Tong University, China

[2]University of Massachusetts at Boston

[3]University of South Florida

October, 2016

Background

- Smart mobile devices are everywhere

- The rise of mobile payment

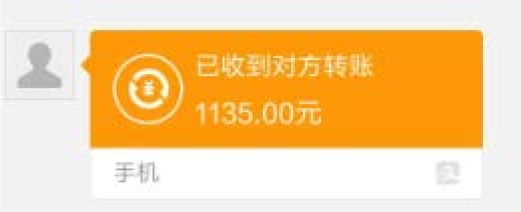
Alipay

WeChat

Bank APP

Online Mobile Payment

Quick Pay



已收到对方转账

1135.00元

手机

Money transfer

Online payment

Alipay

In 2015

900 million users

100 million transactions per day

1 trillion dollars transactions

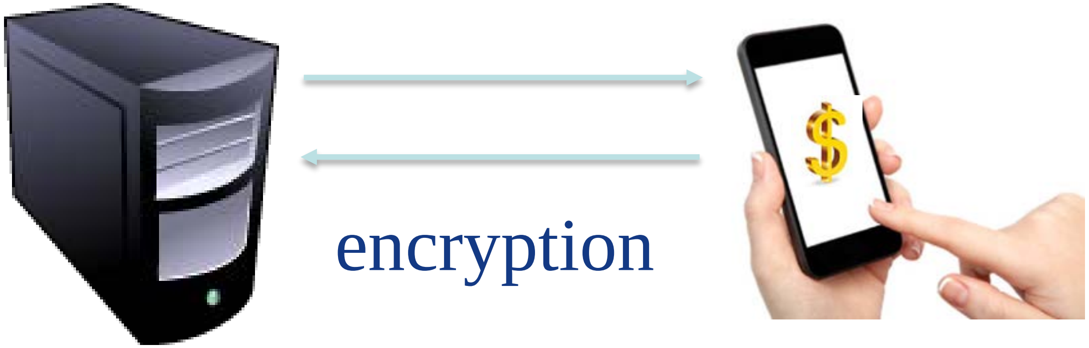
Payment Protections

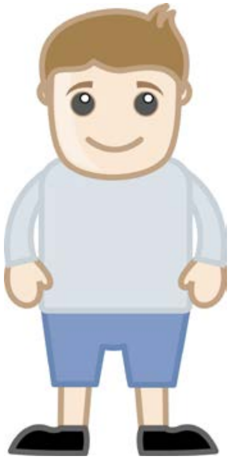
Protections of mobile payment security

- Transport protocol: TLS/SSL

The packets payloads are encrypted

- 6-digit Password





Trust

- Limited password attempt times

Payment Protections

Protections of mobile payment security

- Transport protocol: TLS/SSL

The packets payloads are encrypted

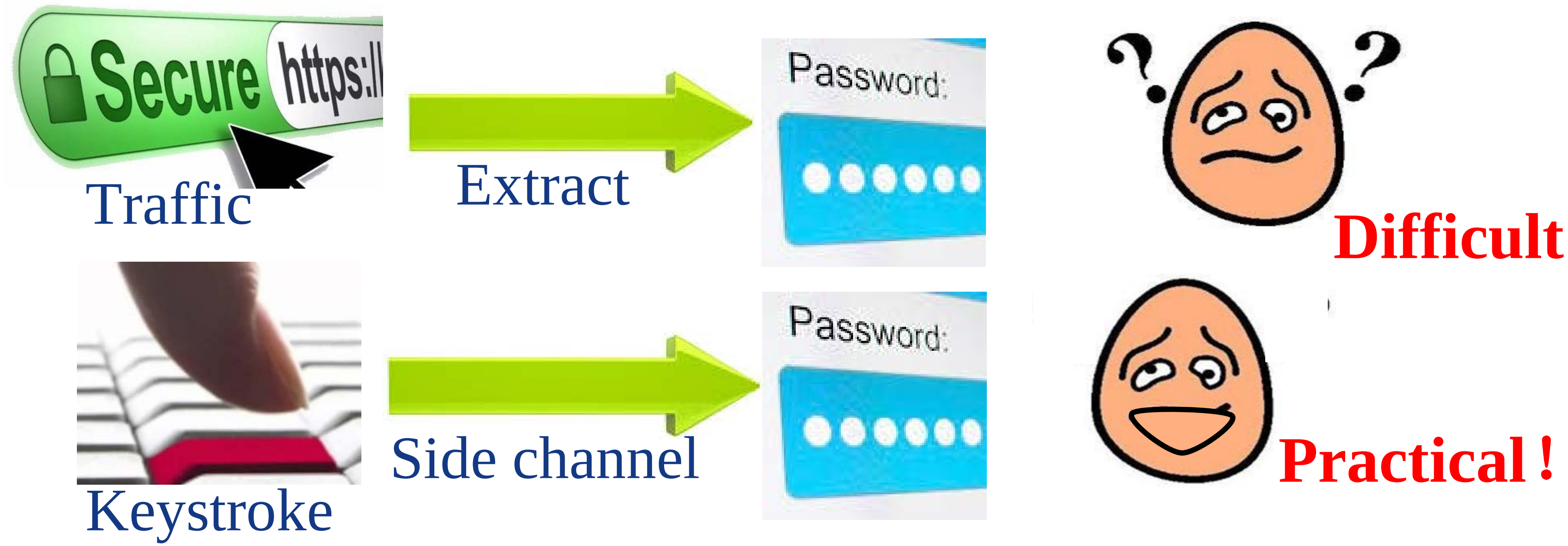
- 6-digit Password

encryption

Danger!

- Limited password attempt times

Password Inference



Traffic

Extract

Difficult

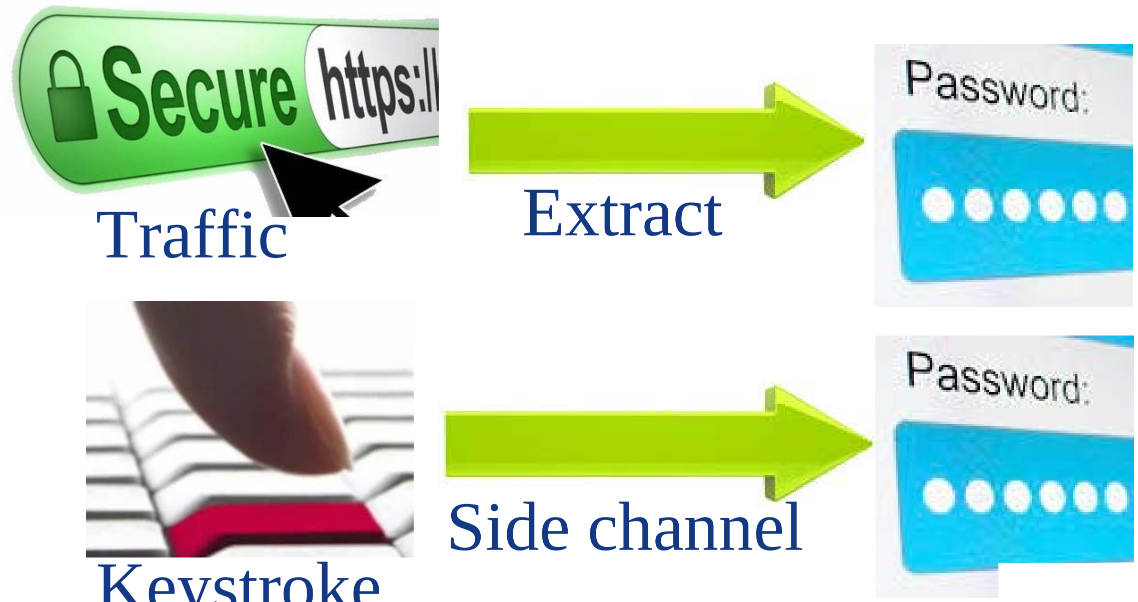
Keystroke

Side channel

Practical !

- Keystroke Inference methods:
  - Accelerometer based method: CCS 2015
  - Acoustic based method: CCS 2014
  - Camera based method: CCS 2014
- Their assumption **cannot** hold in **mobile payment** scenario.

PASSWORD INFERENCE



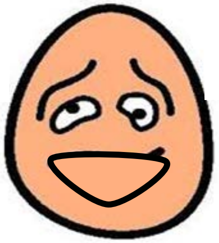
Traffic

Extract

Keystroke

Side channel

Difficult



Practical !

- Keystroke Inference Models:
  - Accelerometer based meth
  - Acoustic based method: C
  - Camera based method: CC
- Their assumption **cannot** hol scenario.

Specifically:

Channel State Information (CSI) from Wi-Fi

Channel State Information

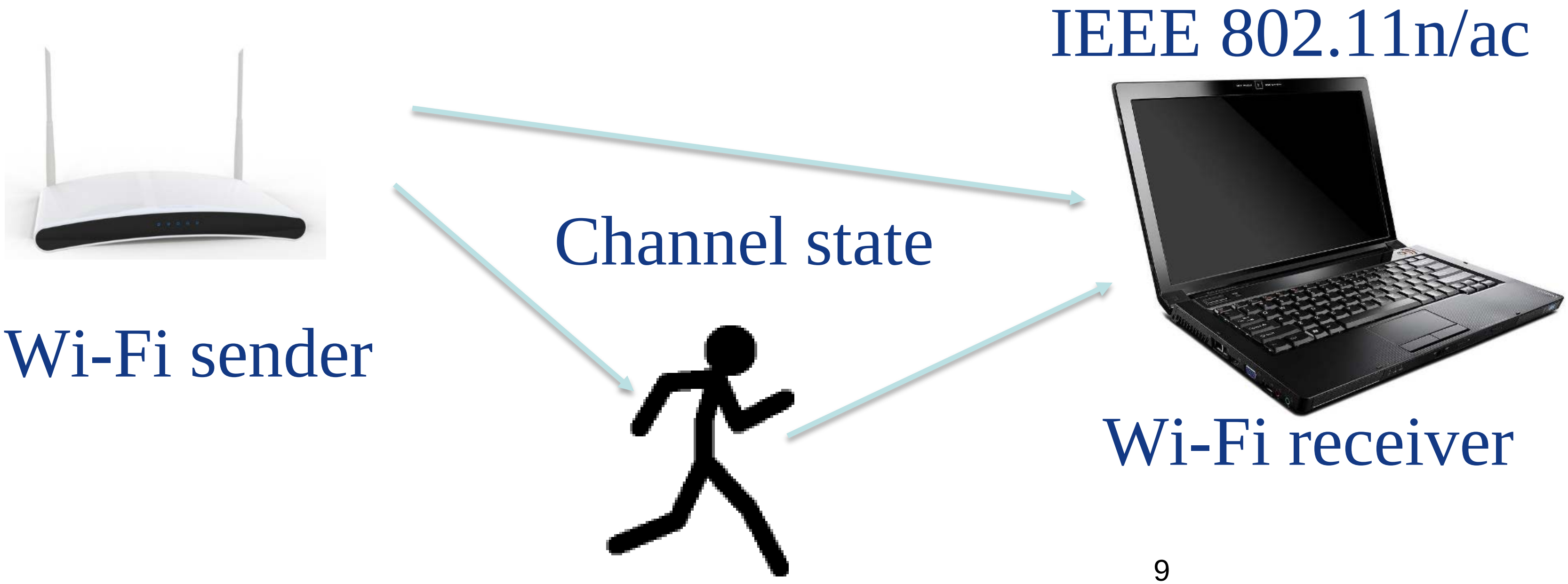
- CSI(Channel State Information)

  CSI was the **channel frequency response** of Wireless signals.

Channel State Information

- CSI(Channel State Information)

  CSI reflects the state of its transmission channel.



IEEE 802.11n/ac

Channel state

Wi-Fi sender

Wi-Fi receiver

Existing Works about CSI Based Recognition

- Centimeters-level Localization

  Chronos D Vasisht, S Kumar, D Kataba (NSDI 2016)

- Person Indentification

  WiWho Y Zeng, P Pathak, P Mohapatra (IPNS 2016)

- Activity Recognition

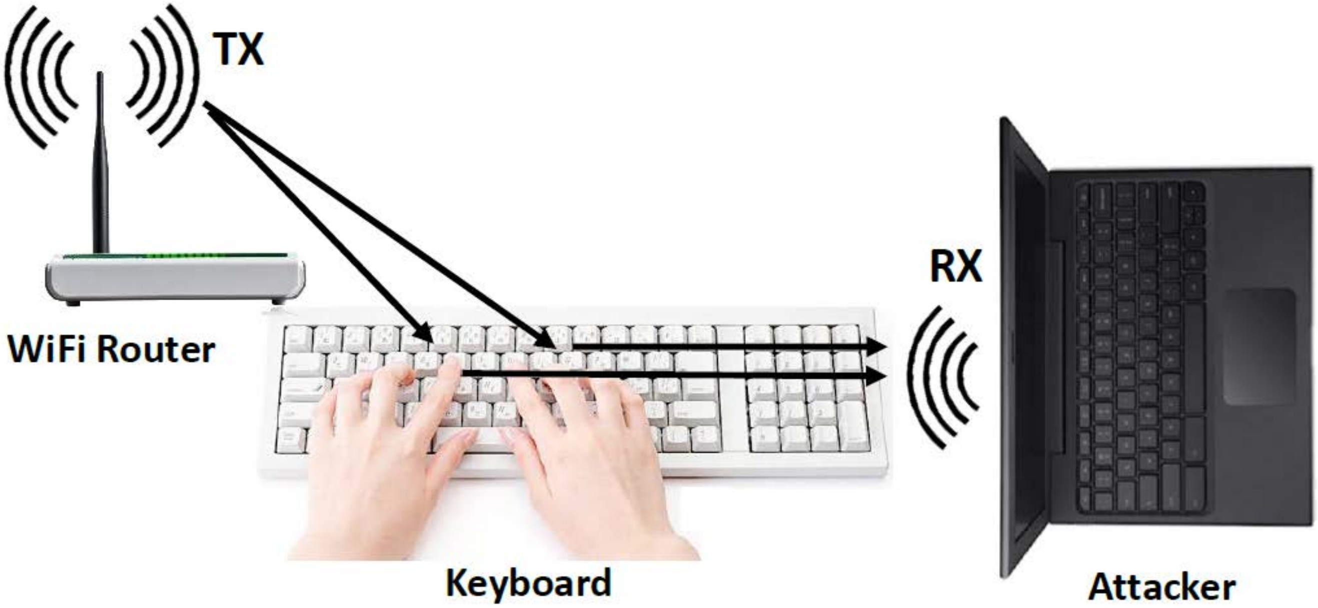
  CARM W Wang, A Liu, M Shahzad, K Ling, S Lu (MobiCom 2015)

- Keystroke Recognition

  WiKey K Ali, A Liu, W Wang, M Shahzad (MobiCom 2015)

Advantage: device-free, commercial equipment

Existing Works about CSI Based Recognition



- Keystroke Recognition

WiKey K Ali, A Liu, W Wang, M Shahzad (MobiCom 2015)

Advantage: device-free, commercial equipment

Existing Works about CSI Based Recognition

- Centimeters-level Localization

  Chronos D Vasisht, S Kumar, D Katabi (NSDI 2016)

- Person I

  WiW

- Activity

  CARI g

  (MobiCom 2015)

- Keystroke Recognit n

  WiKey K Ali, A Liu, W Wang, M Shahzad (MobiCom 2015)

Can existing works be applied to infer payment passwords in mobile devices?

Advantage: device-free, commercial equipment

Existing Works about CSI Based Recognition

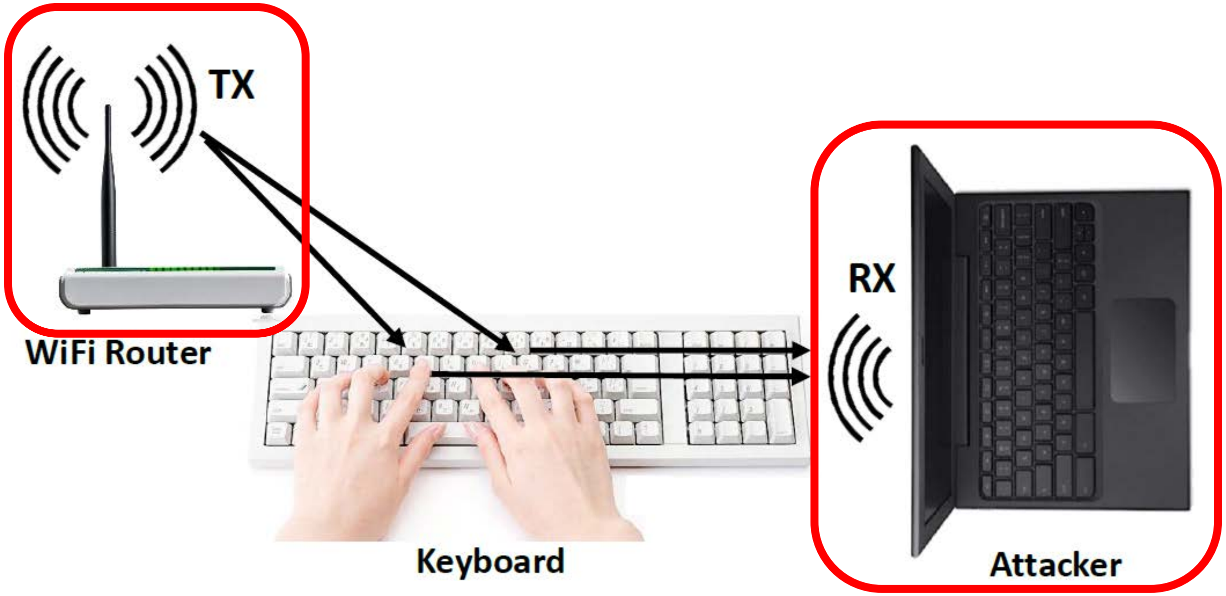
- Centimeters-level Localization

These works have the following shortcomings:
1 Need a sender and receiver Wi-Fi devices
2 Just recognize input, but have no idea what the input is.

CARM W War

- Keystroke Recogniti

WiKey K Ali, A



Existing Works about CSI Based Recognition

- Centimeters-level Localization

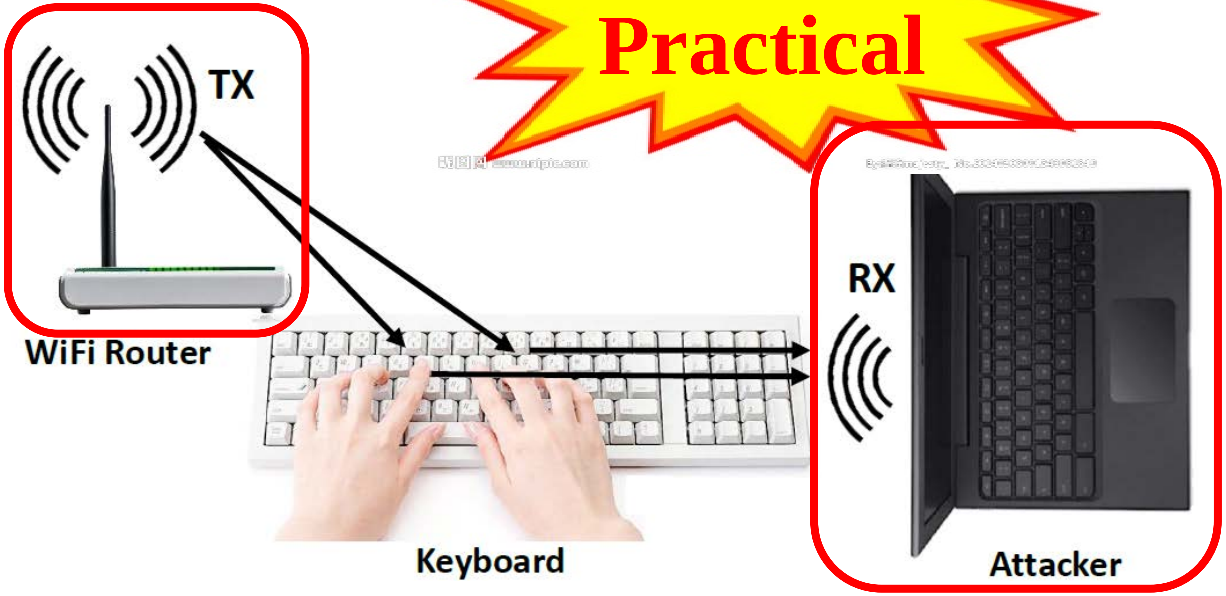
These works have the following shortcomings:
1 Need a sender and receiver Wi-Fi devices
2 Just recognize input, but have no id

Not Practical

CARM W War

- Keystroke Recogniti

WiKey K Ali, A



Our Design -- WindTalker

WindTalker, a novel keystroke inference framework towards Smart Phones through WiFi Channel State Information(CSI).

Feature:

- One device to attack - no requirement of victim locating between two WiFi devices;
- Identifying the sensitive input time window (e.g., password input) by considering the SSL traffic and CSI flow together;
- Successfully attack AliPay, the most popular mobile payment system in the world, on several smart phones.

OUTLINE

- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

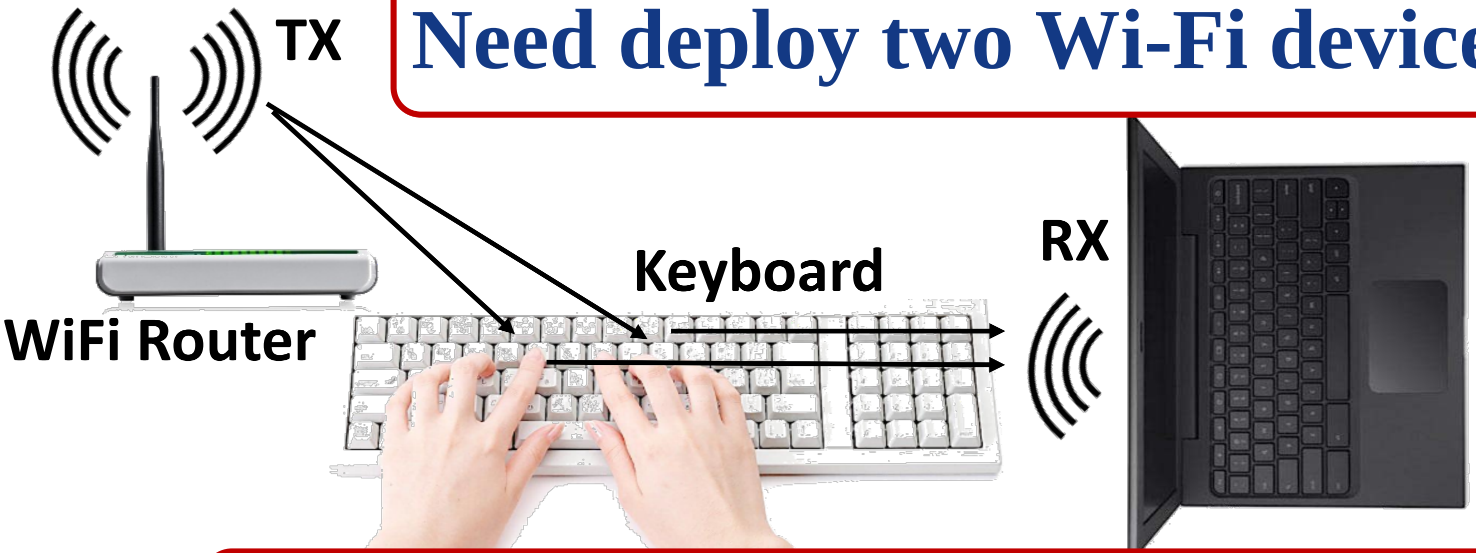
OUTLINE

- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

CSI COLLECTION

- Change CSI collection method to get valid CSI data

Need deploy two Wi-Fi devices

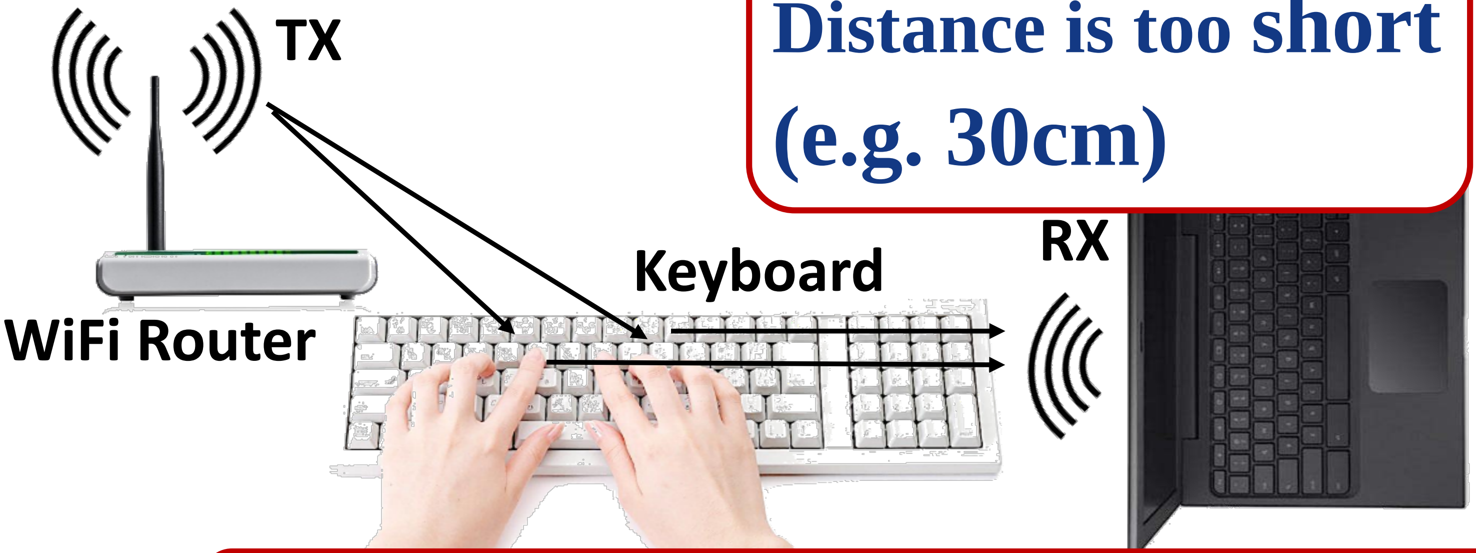


Target locating between two devices

Out-of-band keystroke inference(OKI) model

CSI COLLECTION

- Change CSI collection method to get valid CSI data



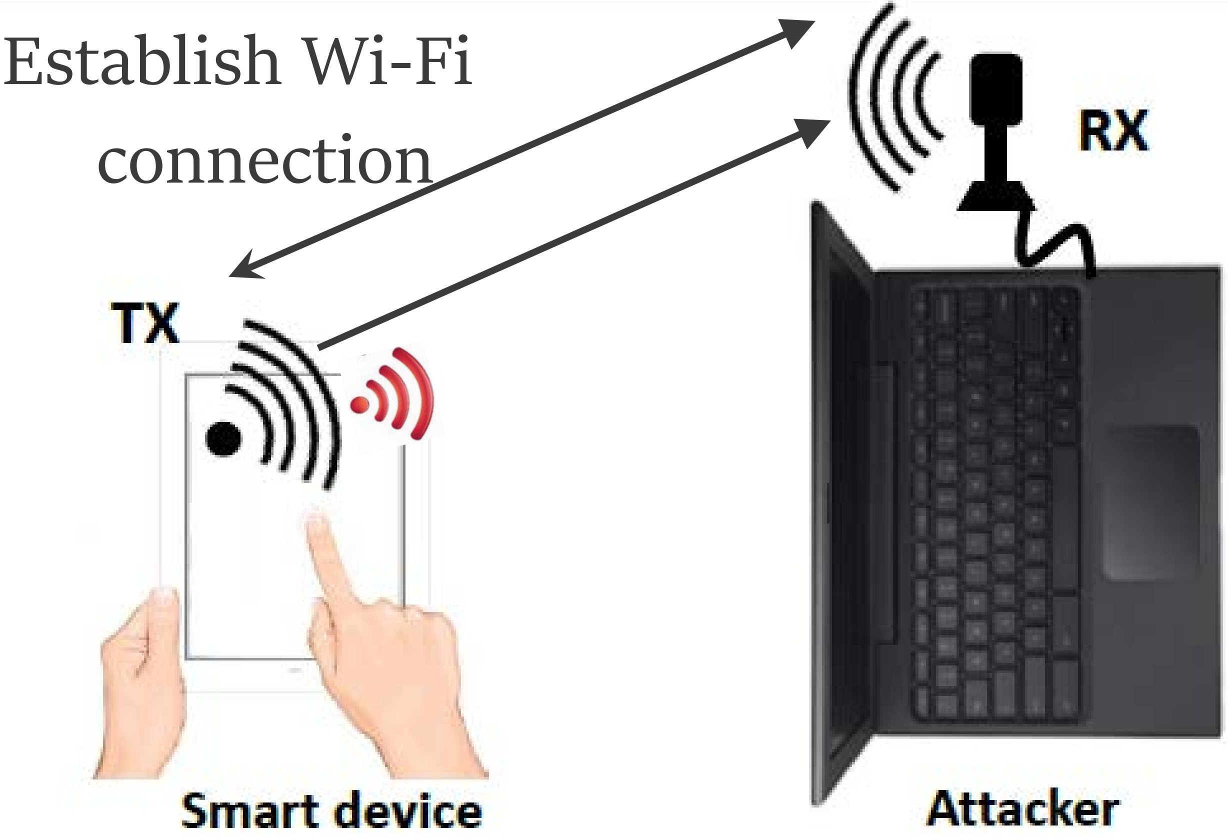
Distance is too short (e.g. 30cm)

Target locating between two devices

Out-of-band keystroke inference(OKI) model

Public WiFi meets CSI – IKI model

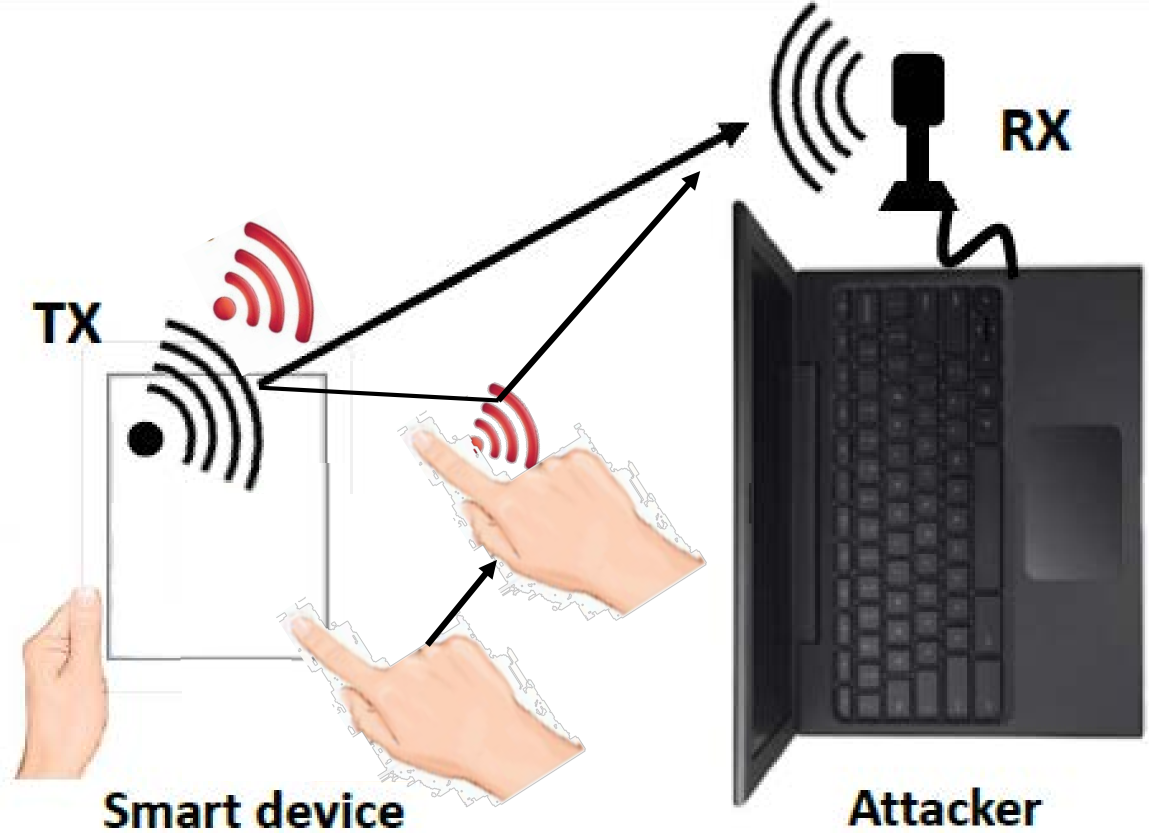
- Change CSI collection method to get valid CSI data



In-band keystroke inference(IKI) model

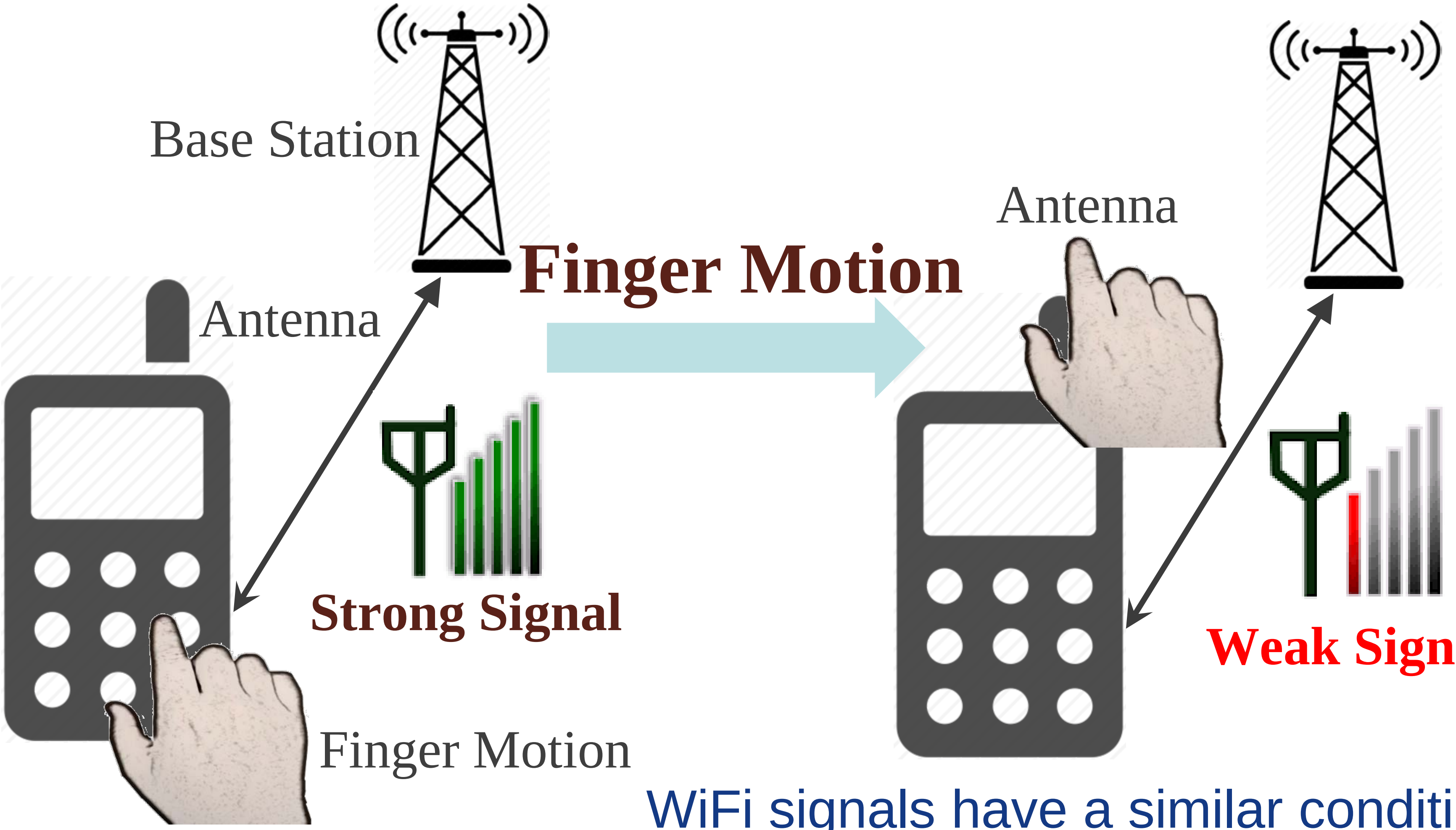
Public WiFi meets CSI – IKI model

- Hand influence– direct influence



CSI - Hand motion

- Factors inference CSI during typing in mobile devices



Base Station

Antenna

Finger Motion

Antenna

Strong Signal

Weak Signal

Finger Motion

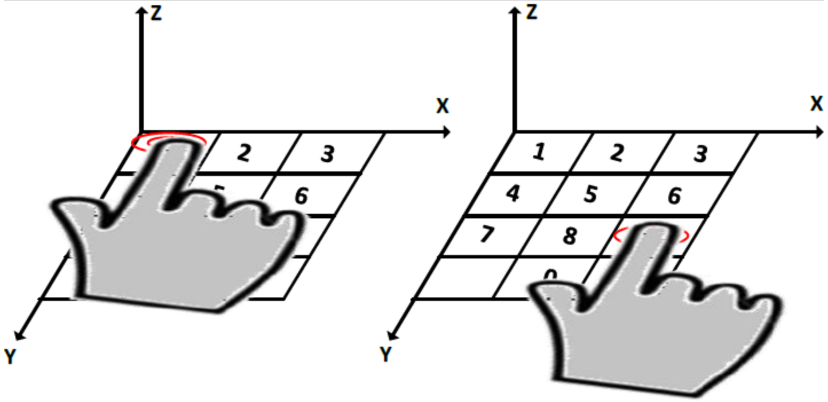
Mobile Phone

WiFi signals have a similar condition.

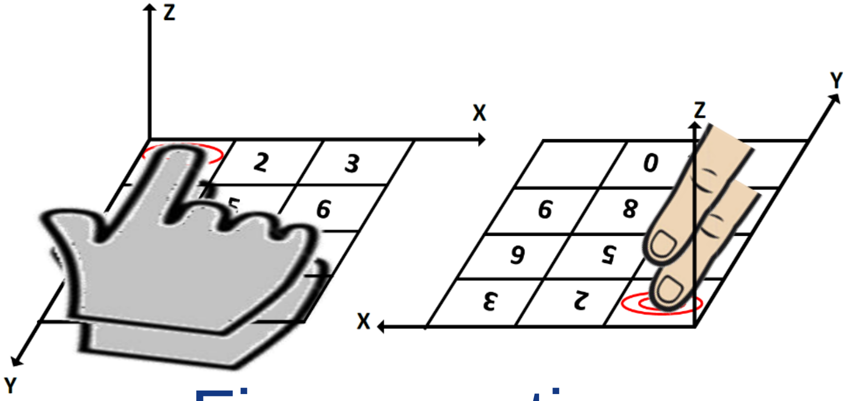
CSI - Hand motion

- Factors inference CSI during typing in mobile devices

Type in soft keyboard



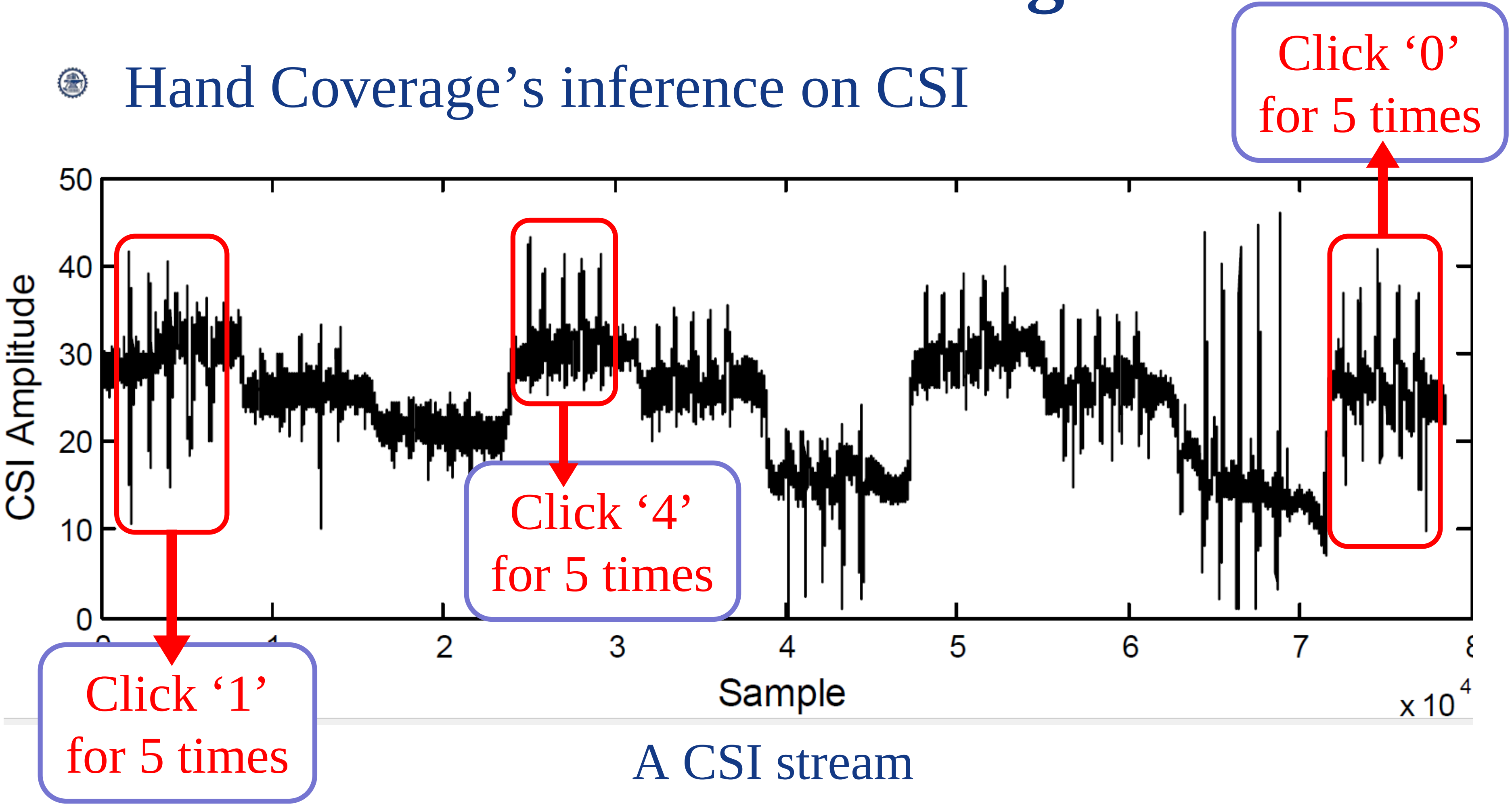
Hand coverage



Finger motion

CSI – Hand coverage

- Hand Coverage's inference on CSI



Click '1' for 5 times

Click '4' for 5 times

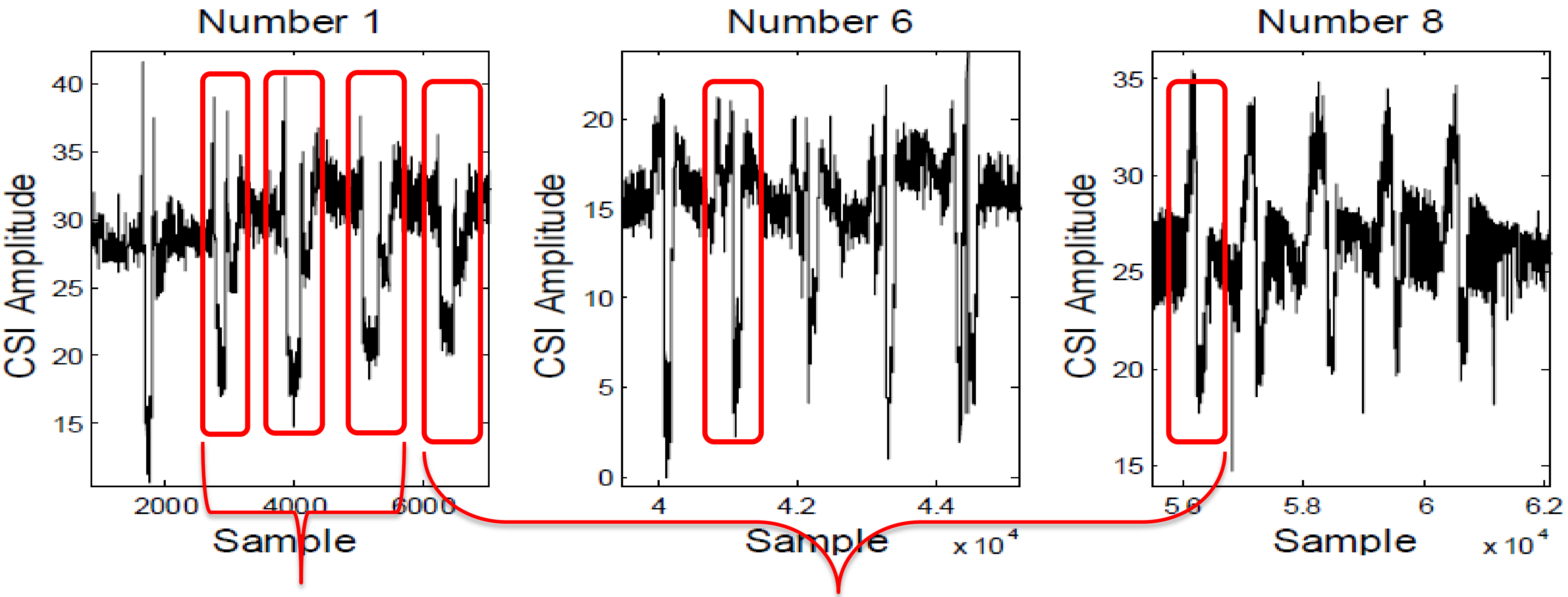
Click '0' for 5 times

A CSI stream

- Continuous press number 1-0 each for 5 times

CSI – Finger motion

- Finger click's inference on CSI– sharp convex



Same numbers

↳ Similarity

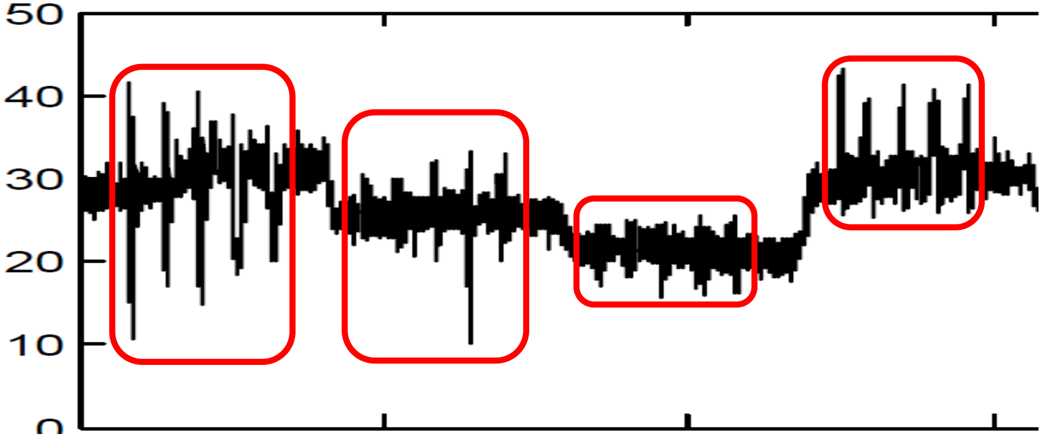
Different numbers

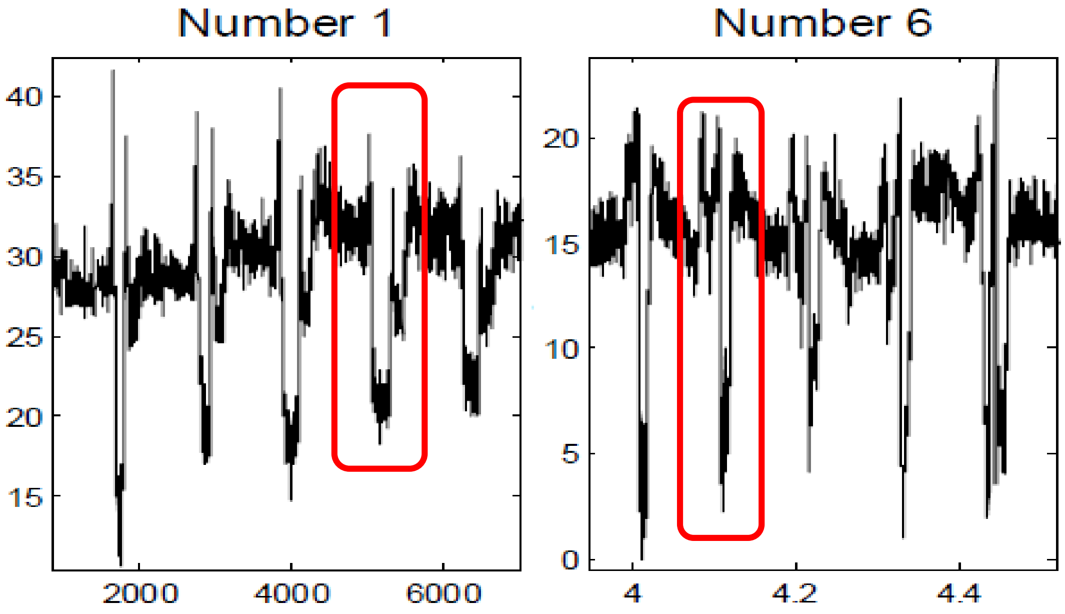
↳ Dissimilarity

Quick click's influence on multi-path propagation

CSI – Finger motion

- Possible



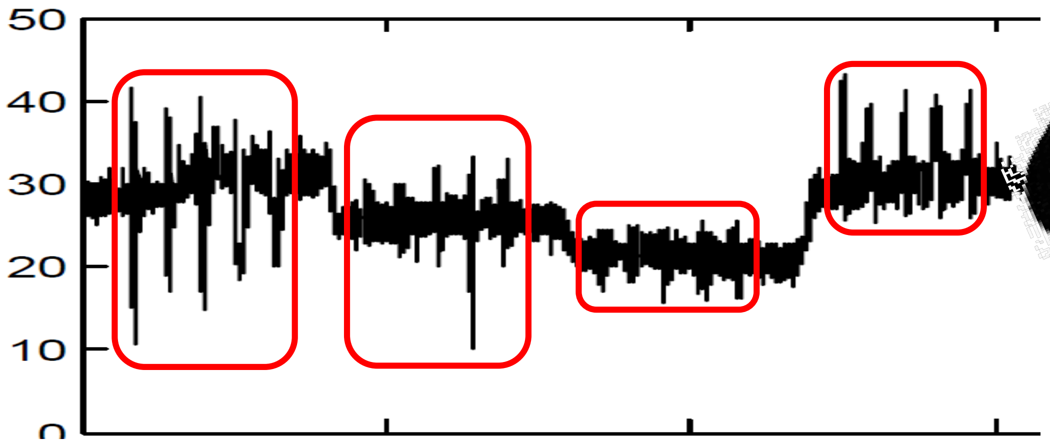


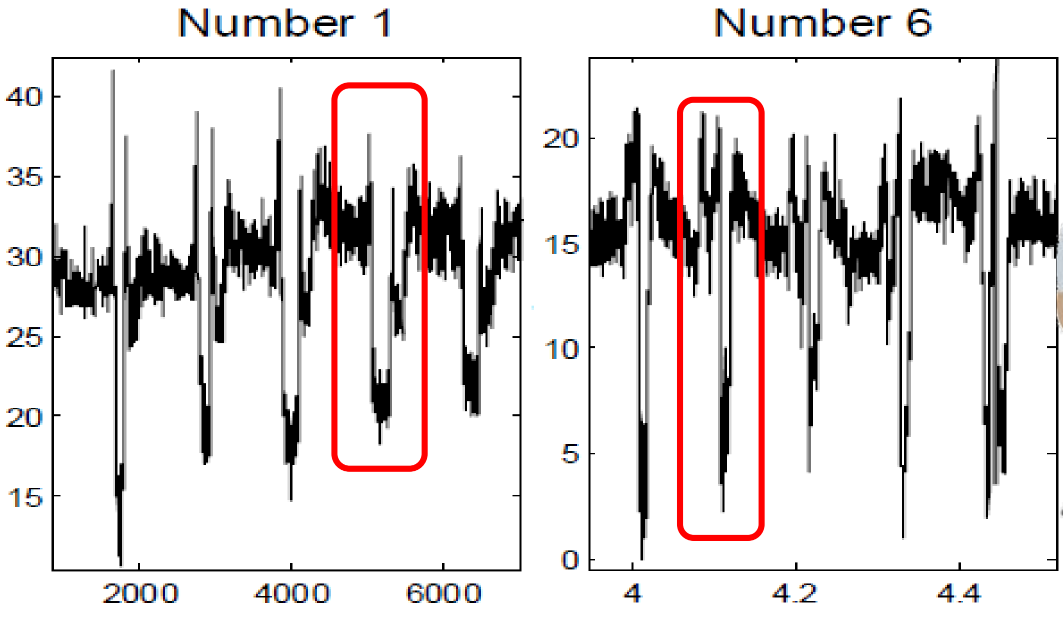
- Possible to **find** finger motion

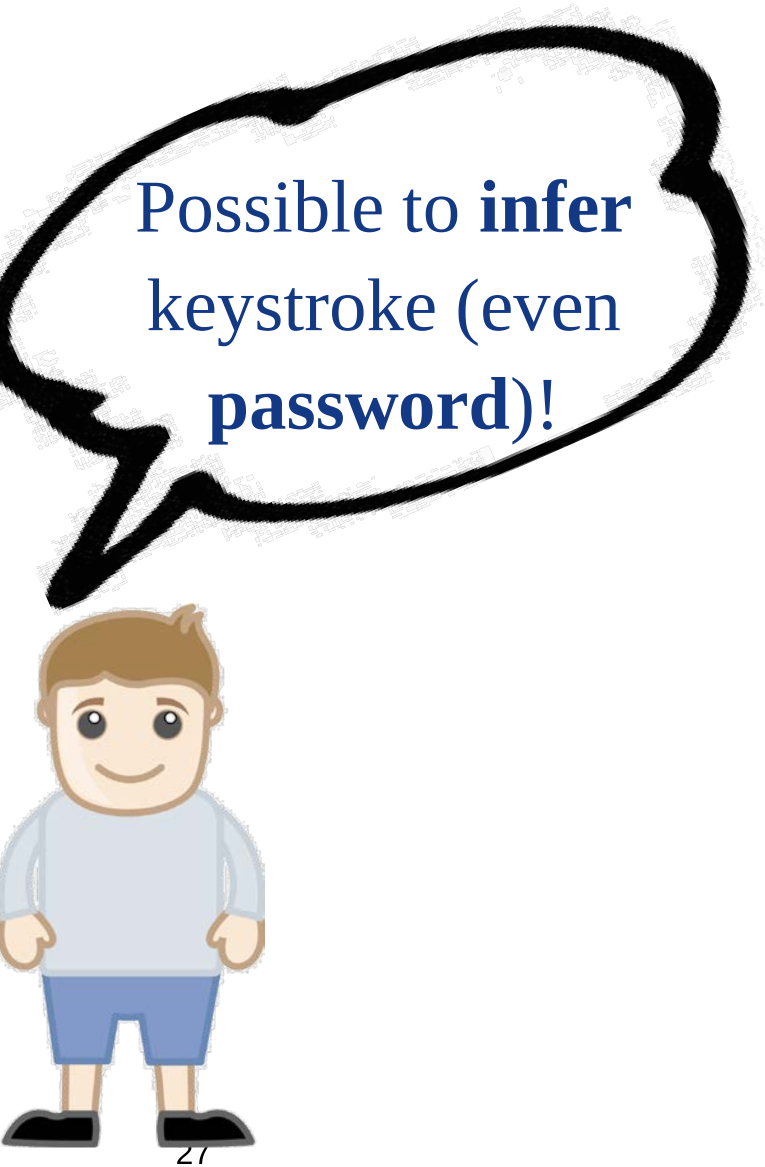
- Possible to **identify** finger motion

CSI – Finger motion

- Possible







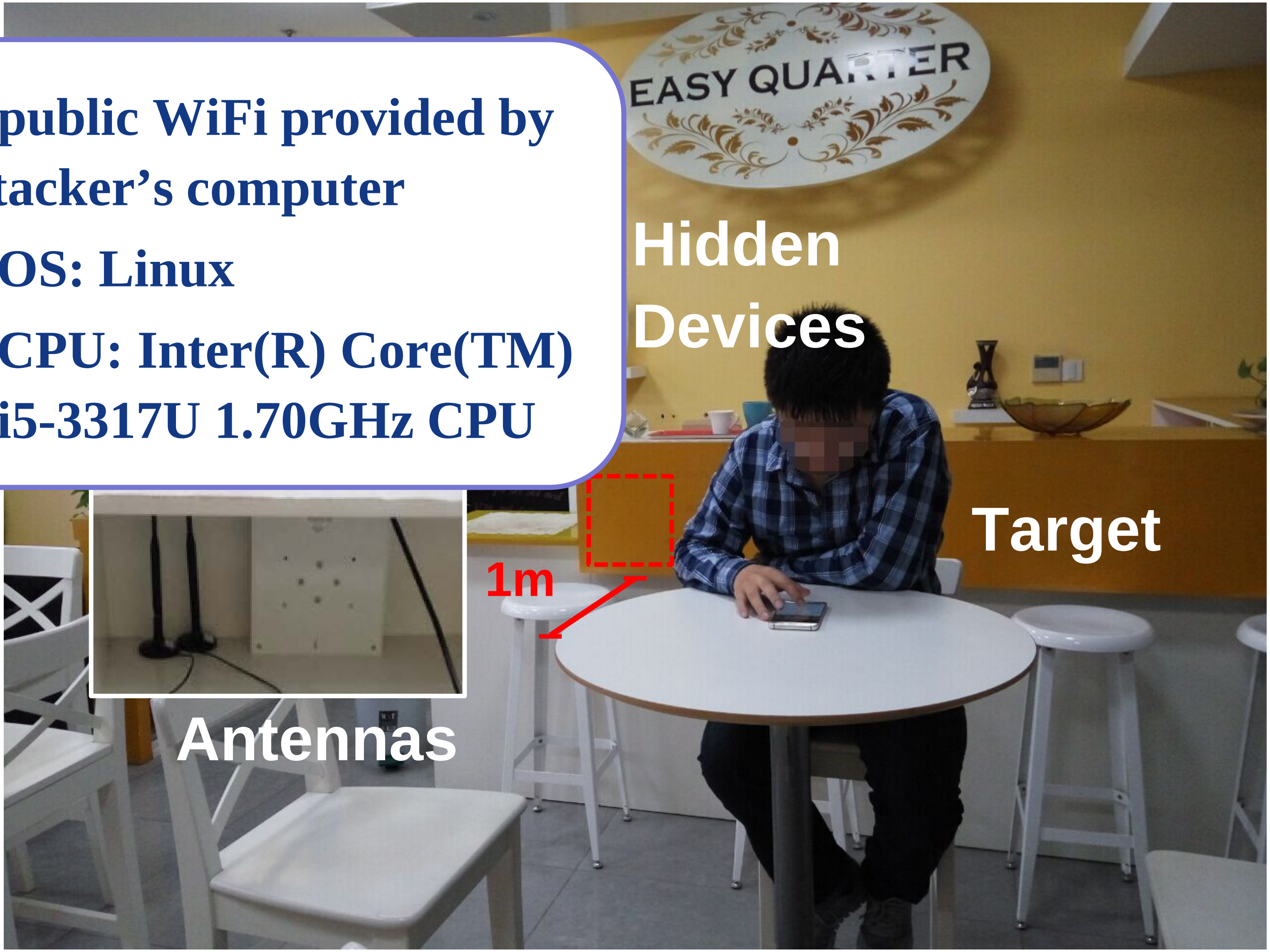
Possible to **infer** keystroke (even **password**)!

OUTLINE

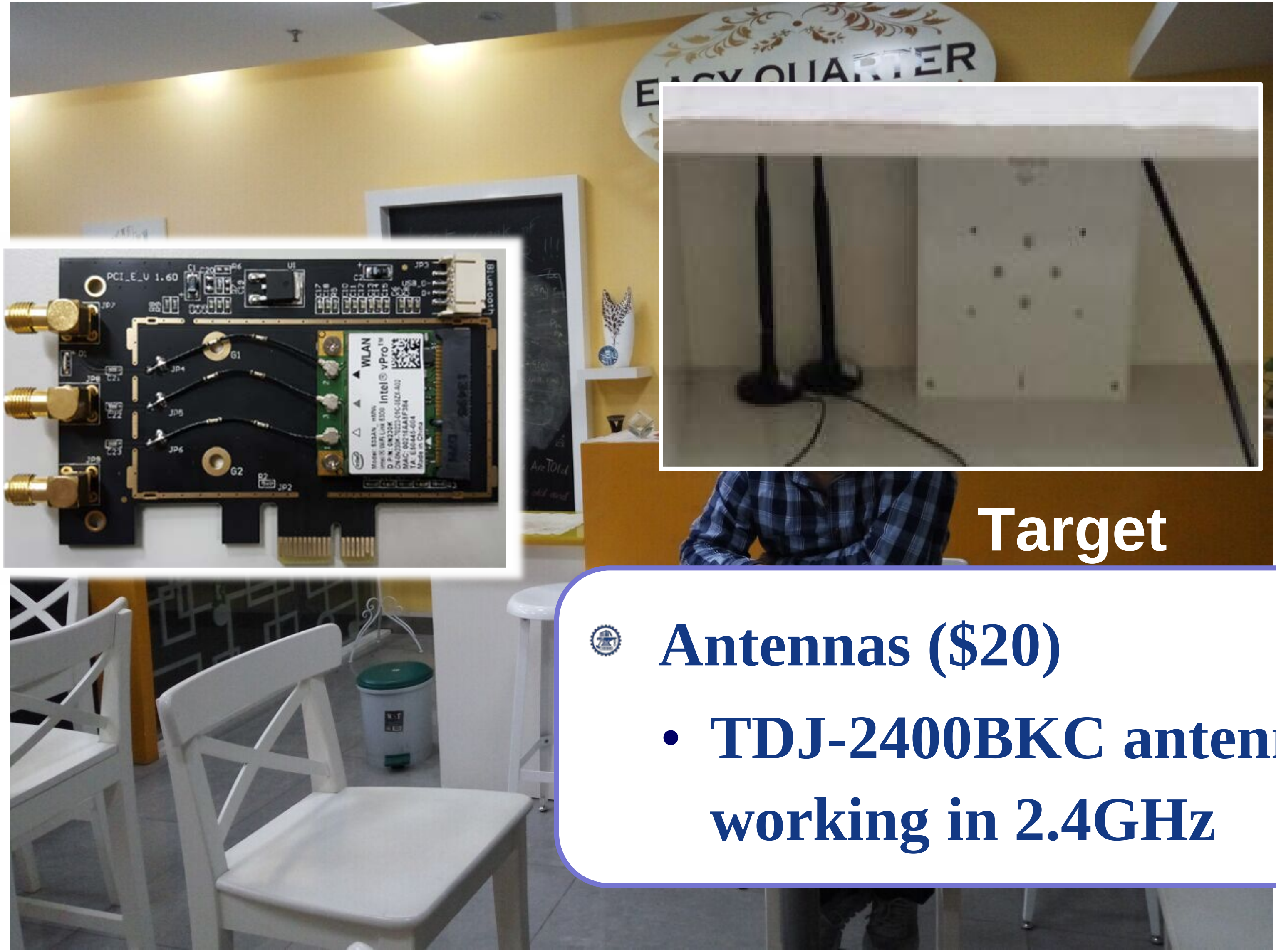
- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

Attack Scenario

- A public WiFi provided by attacker’s computer
  - OS: Linux
  - CPU: Inter(R) Core(TM) i5-3317U 1.70GHz CPU

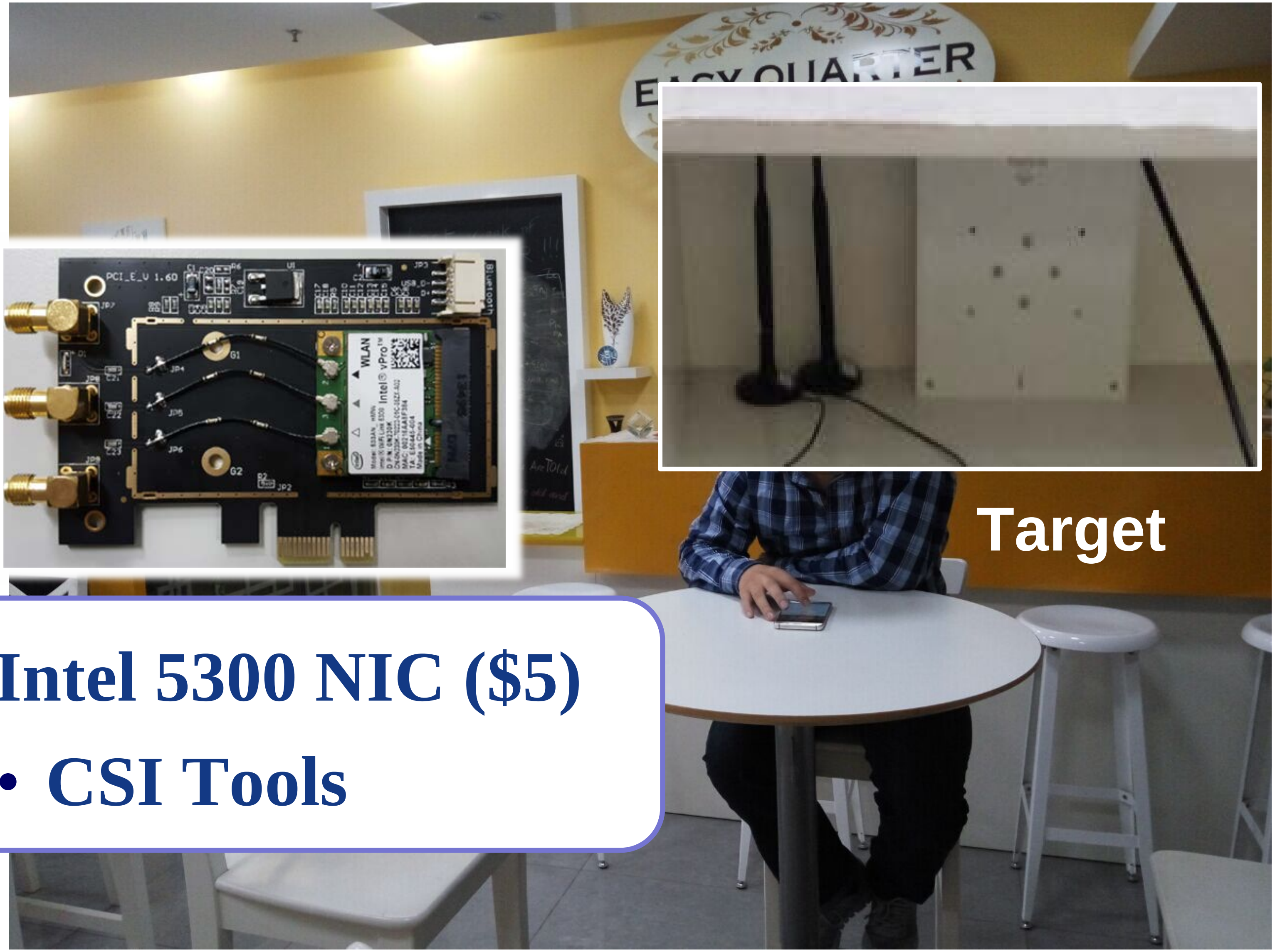


Attack Scenario



- **Antennas ($20)**
  - **TDJ-2400BKC antenna working in 2.4GHz**

Attack Scenario



- Intel 5300 NIC ($5)
  - CSI Tools

OUTLINE

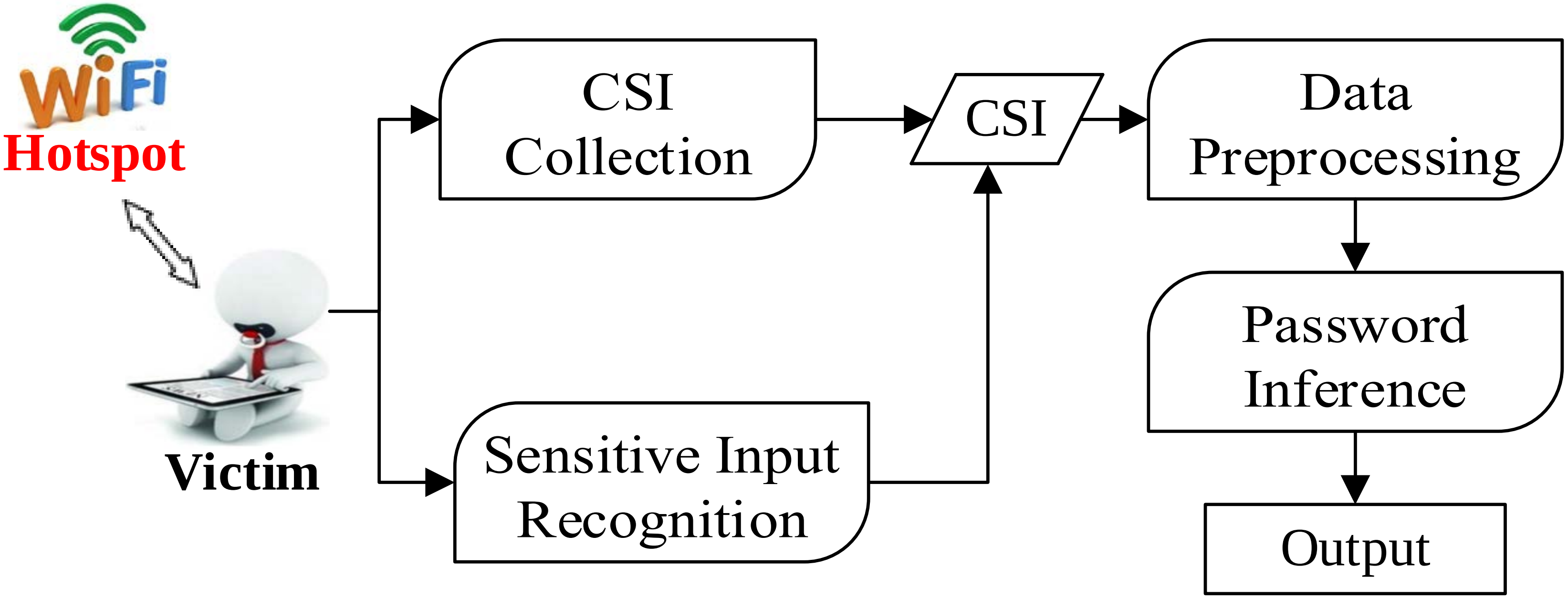
- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

Challenges

- How to enforce victim’s device to be a WiFi sender?
- How to locate CSI segments generated by password input?
- How to reduce noise in raw CSI data?
- How to infer password using CSI?

System Design

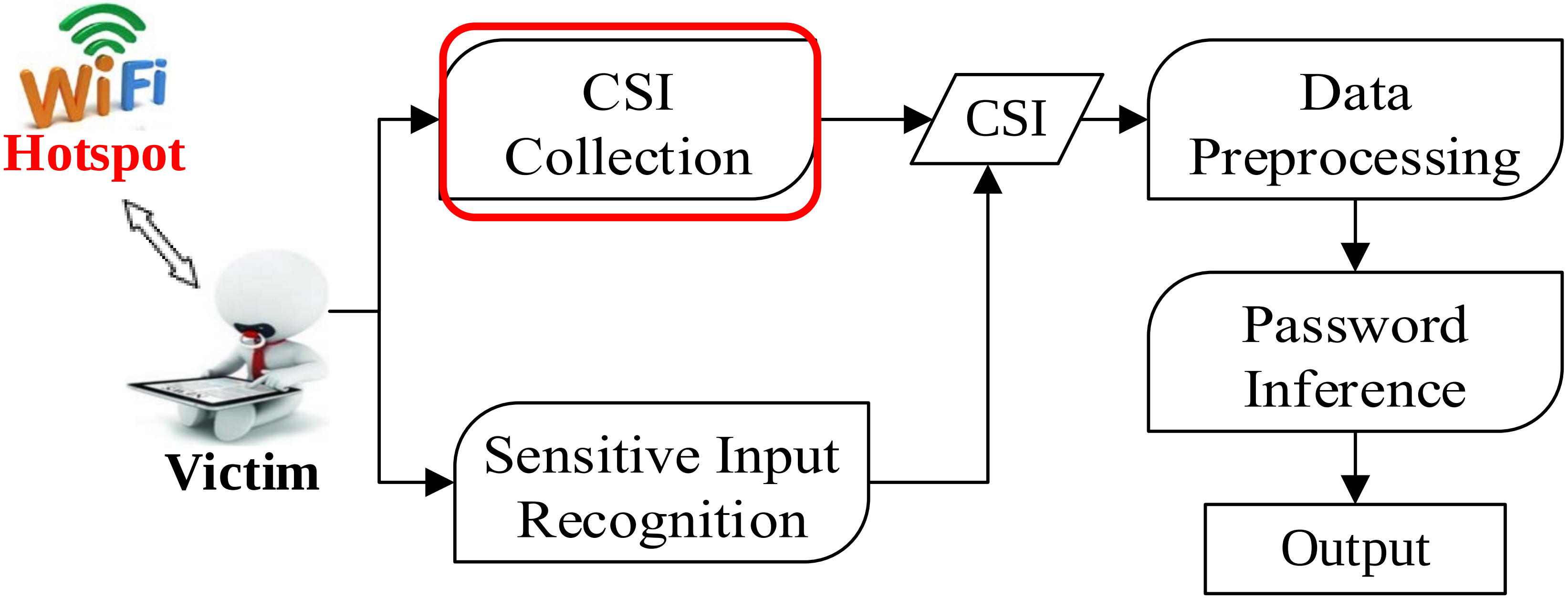
- WindTalker System model
- Four Modules → Four Challenges



WindTalker Schematic

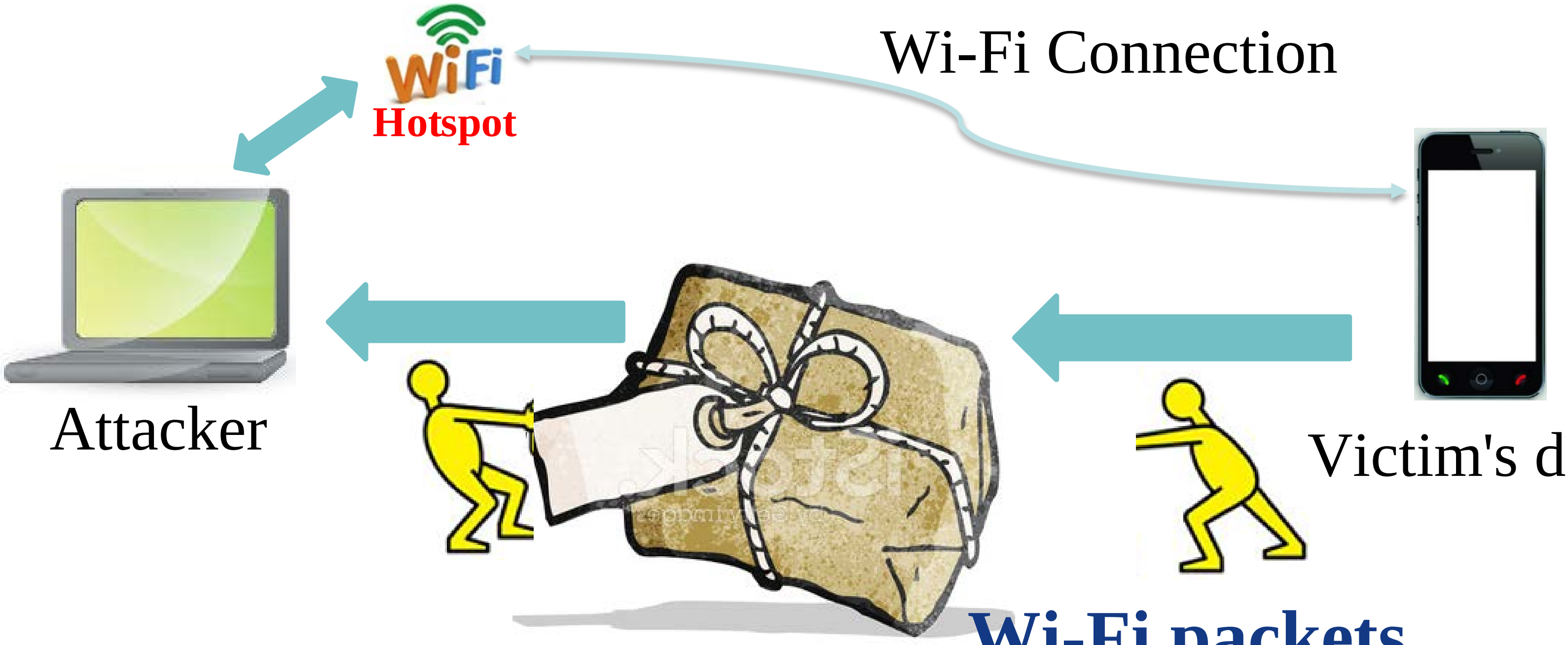
First Challenge

- How to enforce victim's device to be a WiFi sender?
- CSI Collection Module



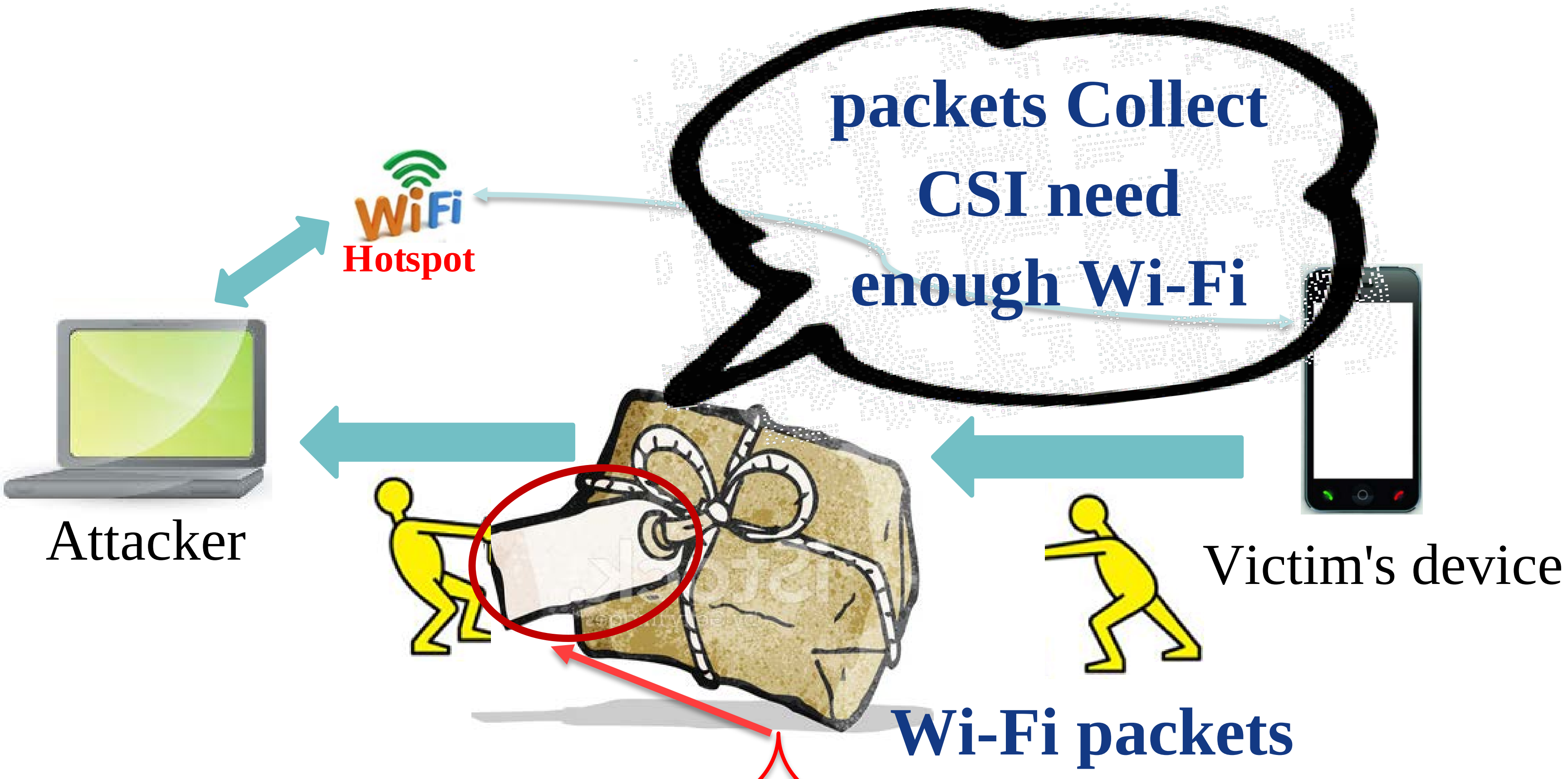
WindTalker Schematic

ICMP based CSI Collection Module



Wi-Fi packets

ICMP based CSI Collection Module



packets Collect
CSI need
enough Wi-Fi

Hotspot

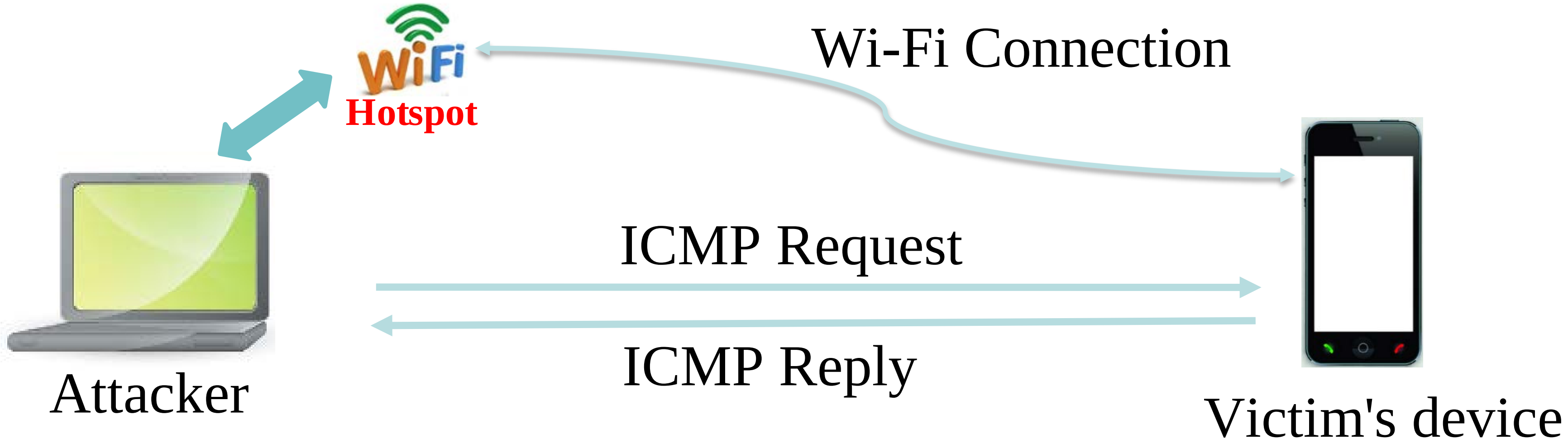
Attacker

Victim's device

Wi-Fi packets

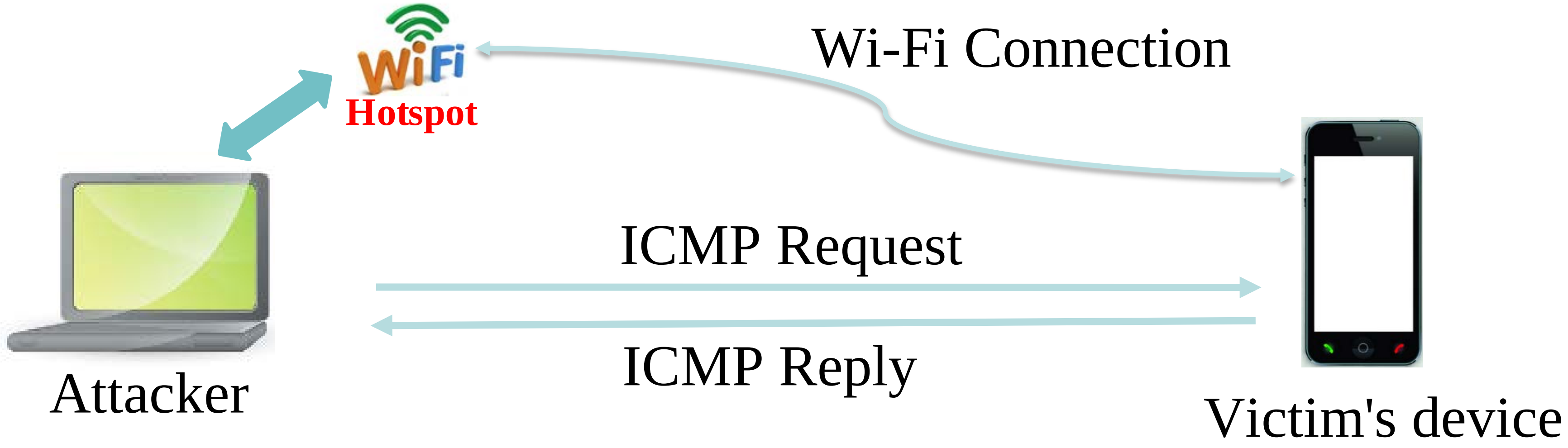
CSI can be extracted from Wi-Fi packets’ preamble

ICMP based CSI Acquirement Module



Attacker sending ICMP request in 800Hz, getting CSI data in 800Hz

ICMP based CSI Acquirement Module

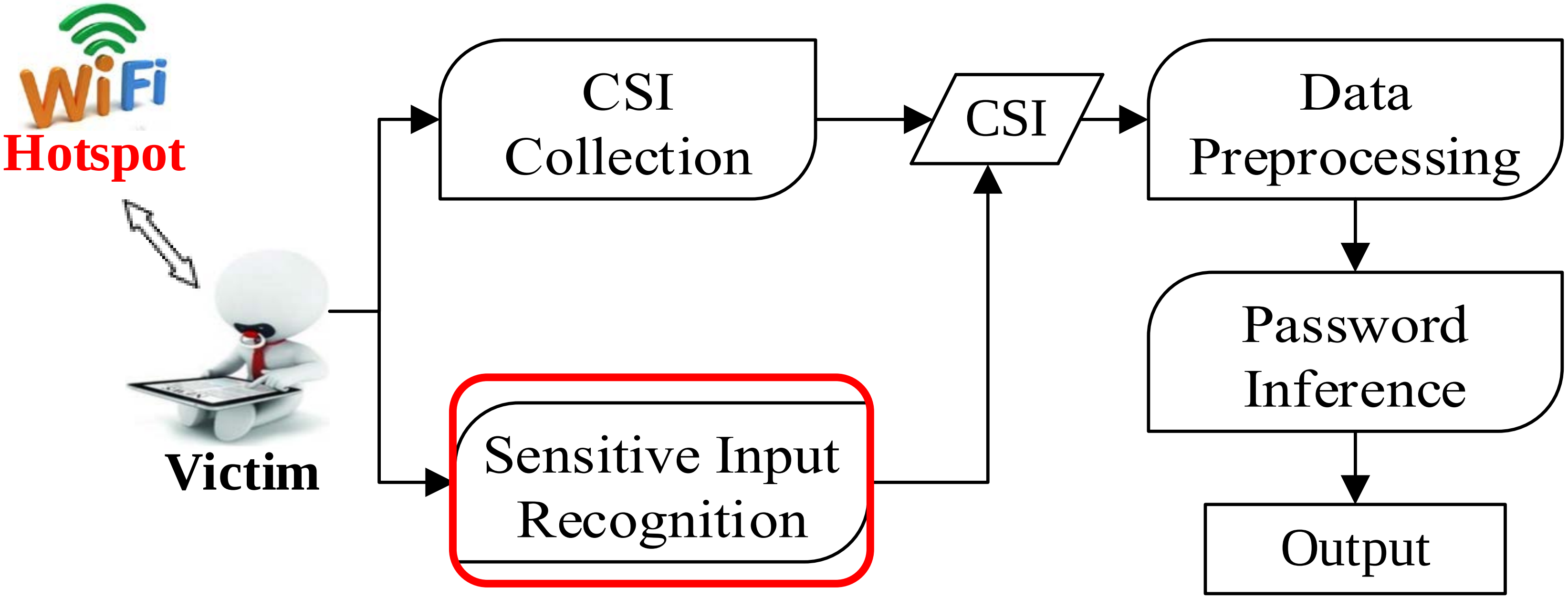


Attacker sending ICMP request in 800Hz, getting CSI data in 800Hz

Can be done without victim’s awareness

Second Challenge

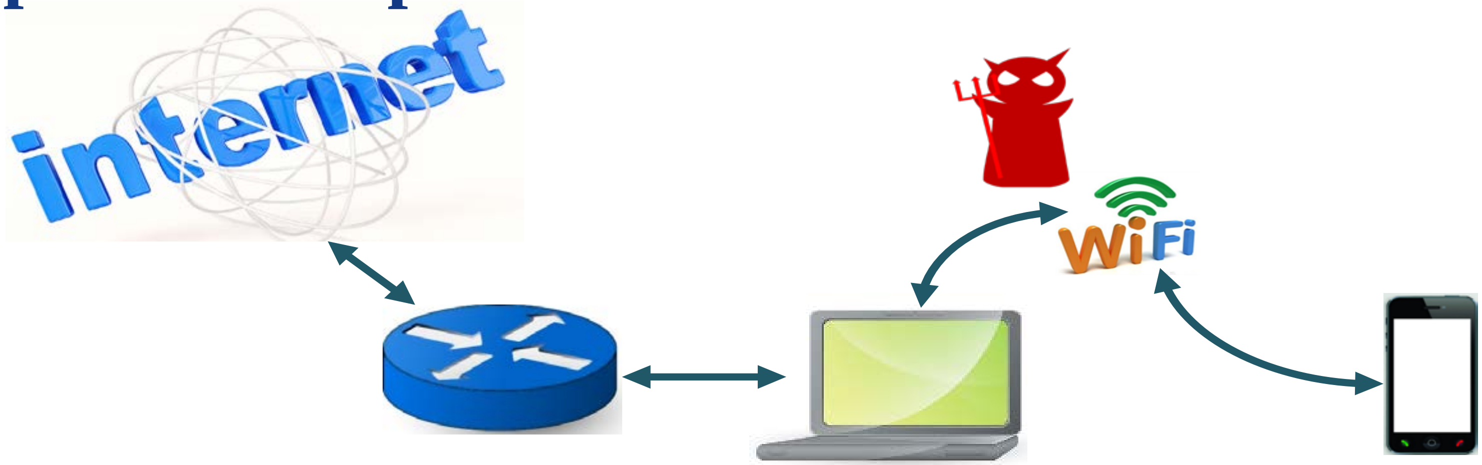
- How to locate CSI segments generated by password input?
- Sensitive Input Module

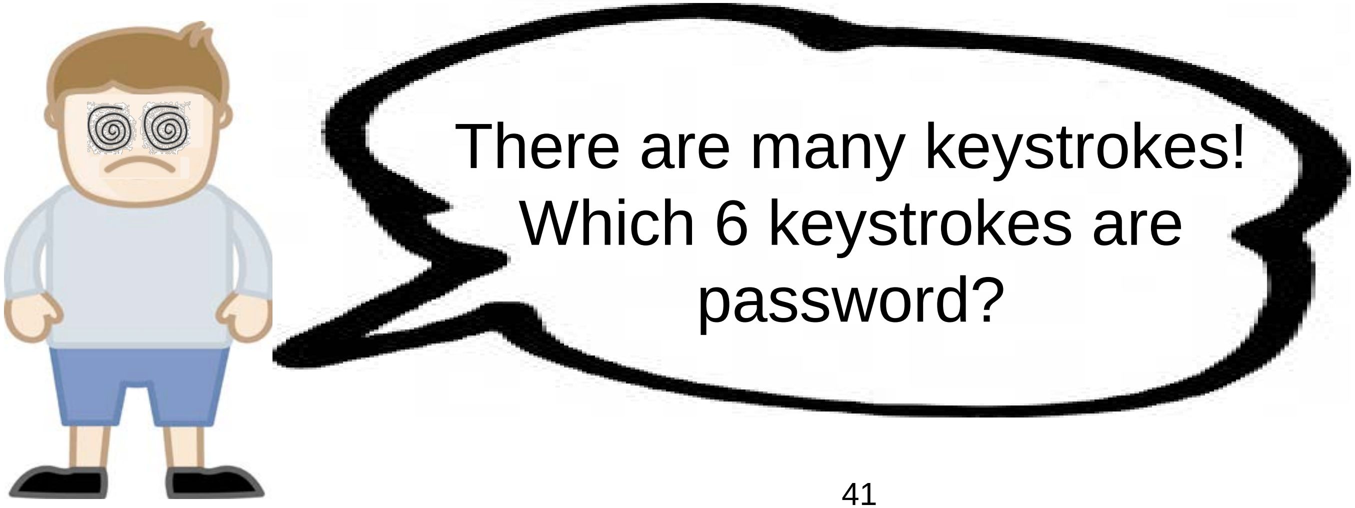


WindTalker Schematic

Sensitive Input Module

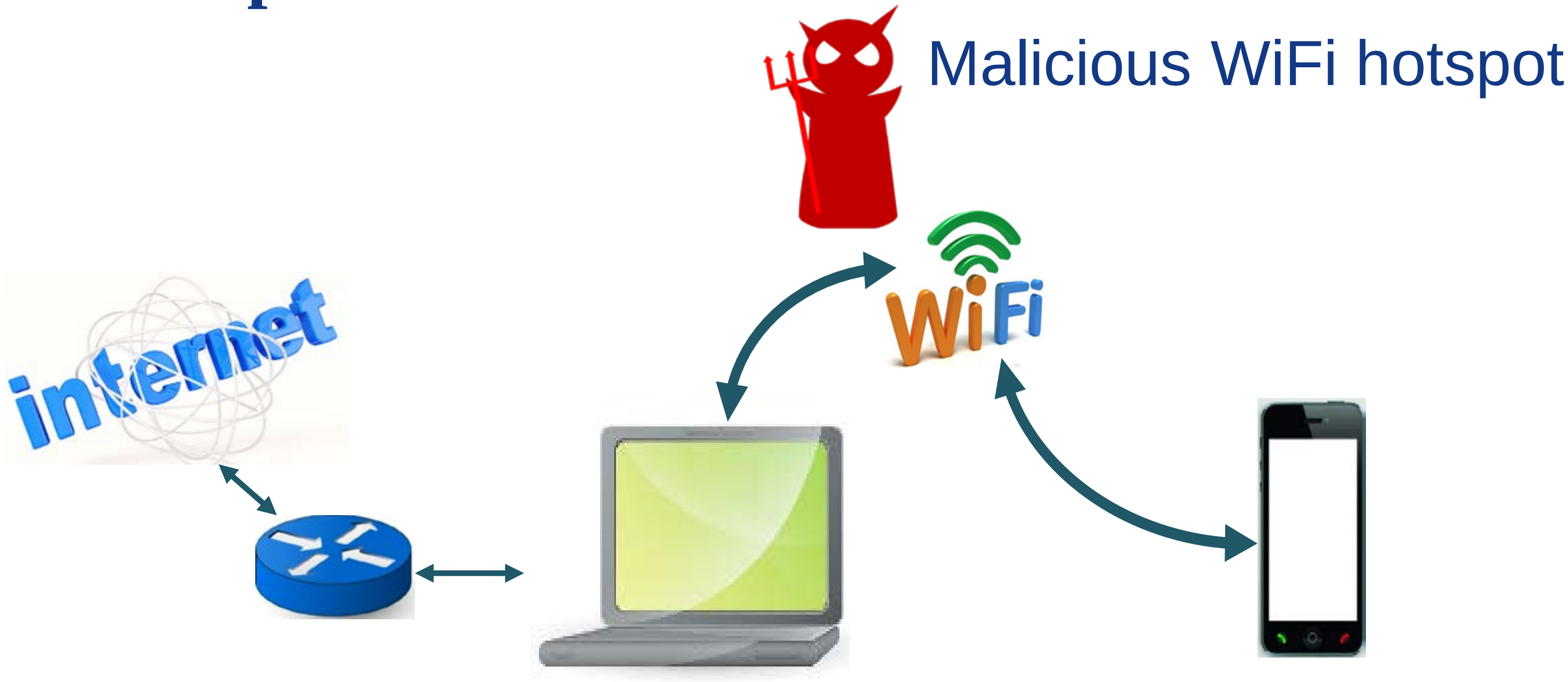
- **How to locate CSI segments generated by password input?**





Sensitive Input Module

- **How to locate CSI segments generated by password input?**



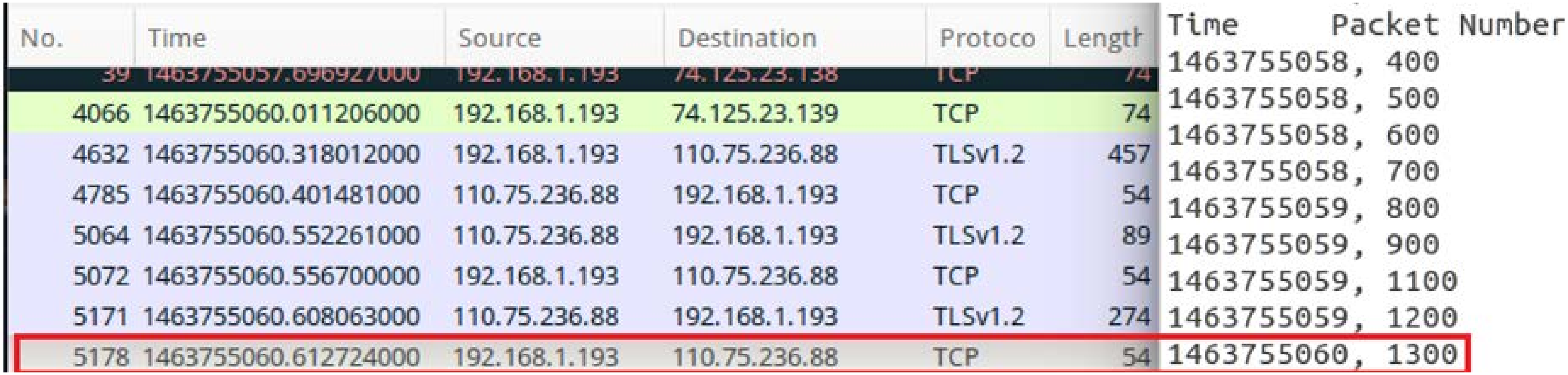
Malicious WiFi hotspot

Make the system more efficient

Sensitive Input Module

- **How to locate CSI segments generated by password input?**

Malicious WiFi hotspot



No.	Time	Source	Destination	Protoco	Length
39	1463755057.696927000	192.168.1.193	74.125.23.138	TCP	74
4066	1463755060.011206000	192.168.1.193	74.125.23.139	TCP	74
4632	1463755060.318012000	192.168.1.193	110.75.236.88	TLSv1.2	457
4785	1463755060.401481000	110.75.236.88	192.168.1.193	TCP	54
5064	1463755060.552261000	110.75.236.88	192.168.1.193	TLSv1.2	89
5072	1463755060.556700000	192.168.1.193	110.75.236.88	TCP	54
5171	1463755060.608063000	110.75.236.88	192.168.1.193	TLSv1.2	274
5178	1463755060.612724000	192.168.1.193	110.75.236.88	TCP	54

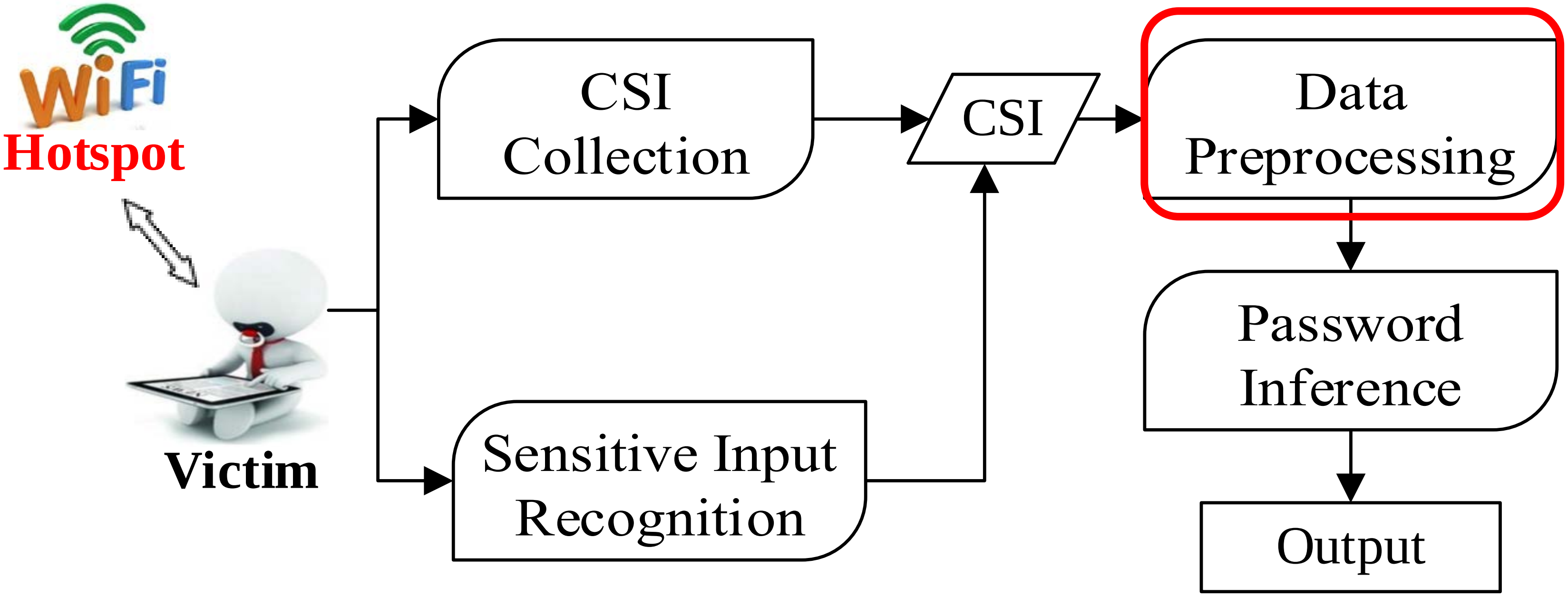
```
Time       Packet Number
1463755058, 400
1463755058, 500
1463755058, 600
1463755058, 700
1463755059, 800
1463755059, 900
1463755059, 1100
1463755059, 1200
1463755060, 1300
```

Construct Sensitive IP Pool

Wireshark

Third Challenge

- How to reduce noise in raw CSI data?
- Data Preprocessing Module

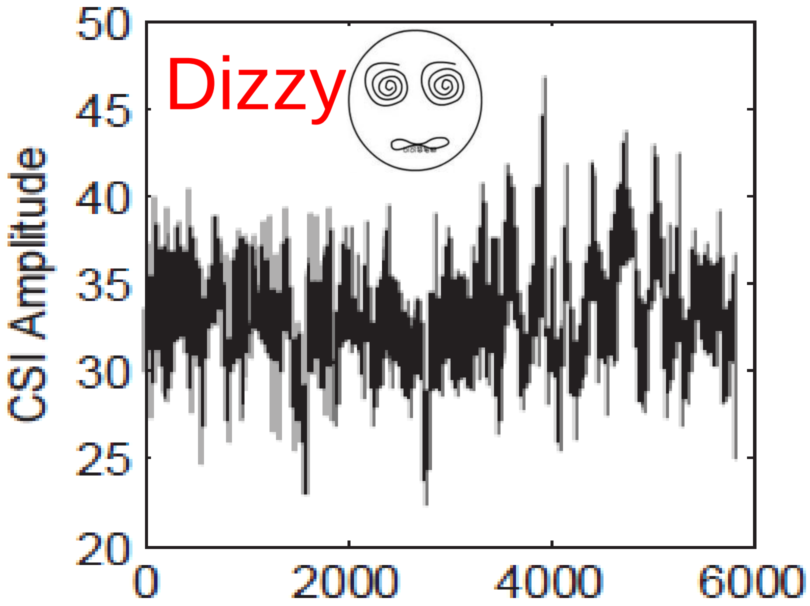


WindTalker Schematic

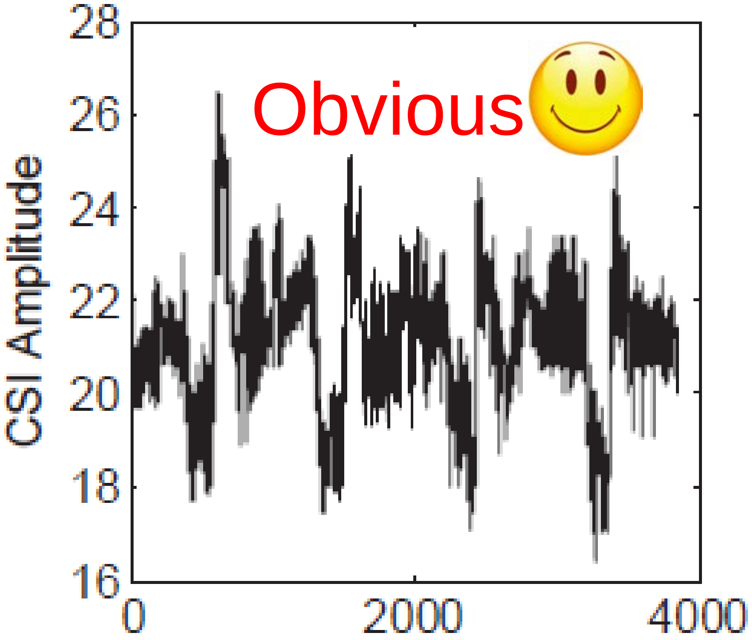
Data Preprocessing Module

- Reducing Noise

  Using Directional Antenna



Using Omni-directional Antenna



Using Directional Antenna

Signal Processing methods

- Reducing Noise

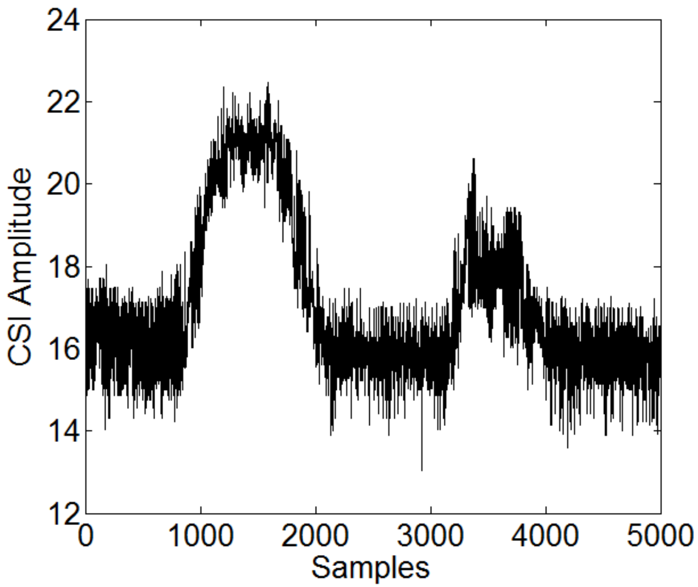
  Low Pass Filtering

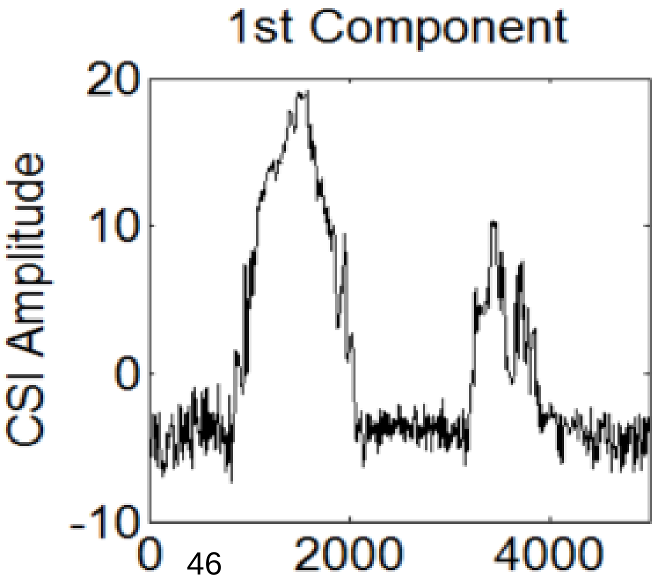
- Dimension Reduction

  Principal Component Analysis (PCA) on subcarriers

  → Select top few projections of CSI data

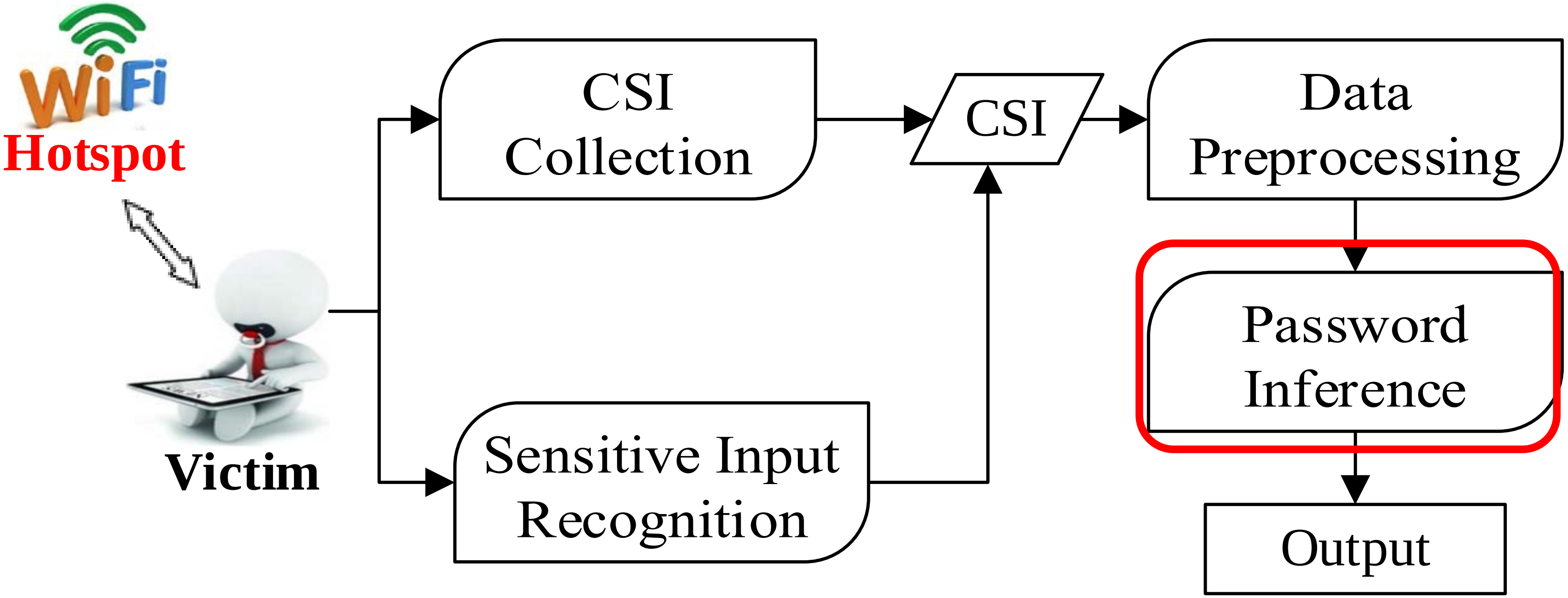
  → Remove the noisy projections of CSI data





Fourth Challenge

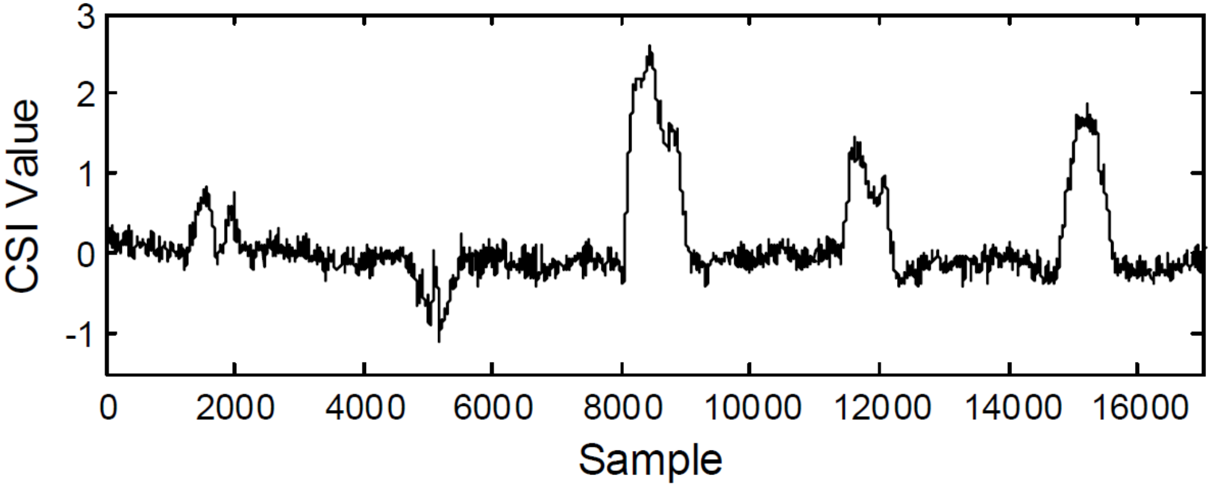
- How to infer password using CSI?
- Data Preprocessing Module



WindTalker Schematic

Password Inference Module

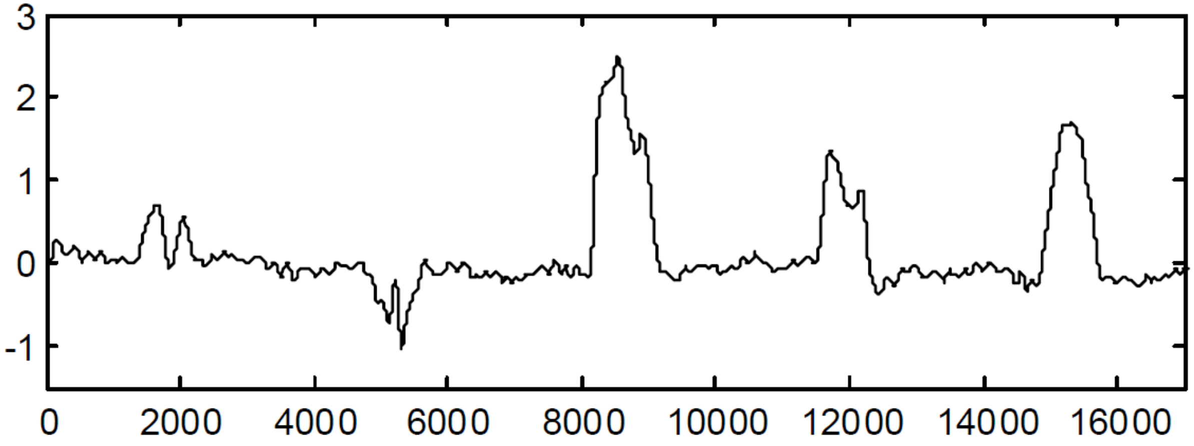
- Keystroke Extraction



Original Data

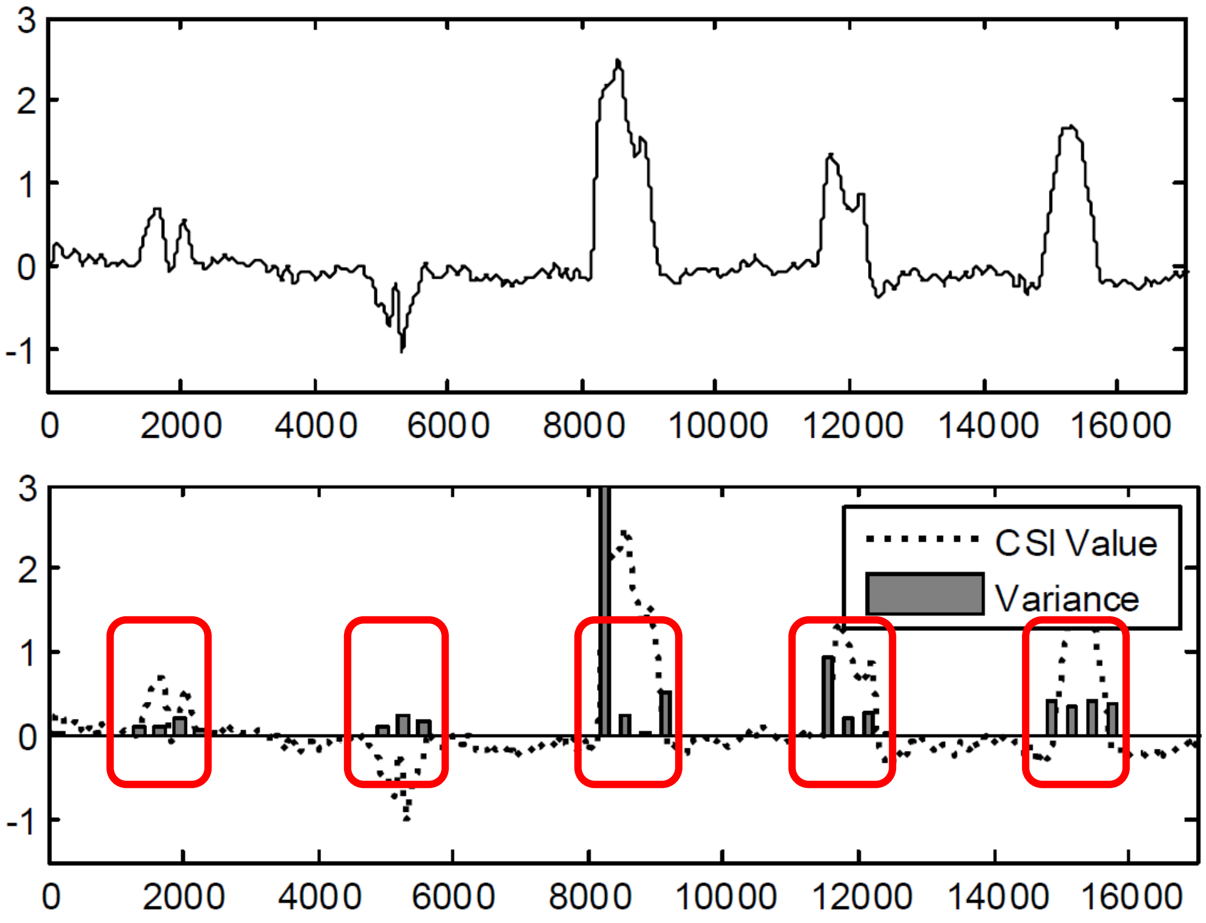
Low-pass Filter

Smooth Data



Password Inference Module

- Keystroke Extraction



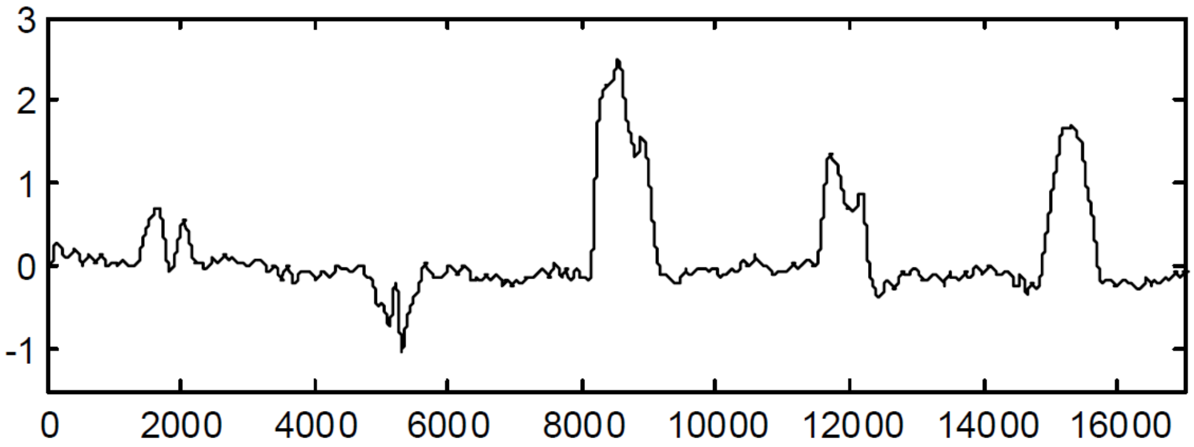
Smooth Data

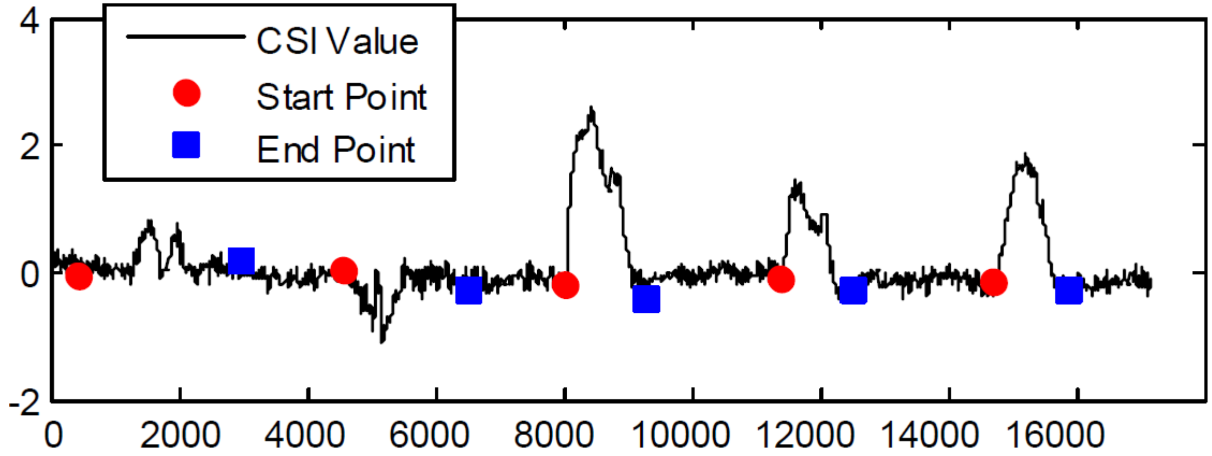
Variance

Choose Segments

Password Inference Module

- Keystroke Extraction





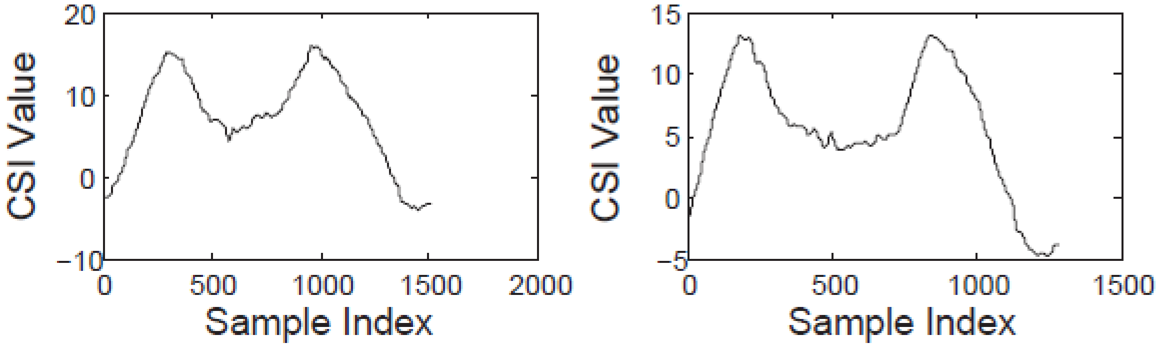
Smooth Data

Variance

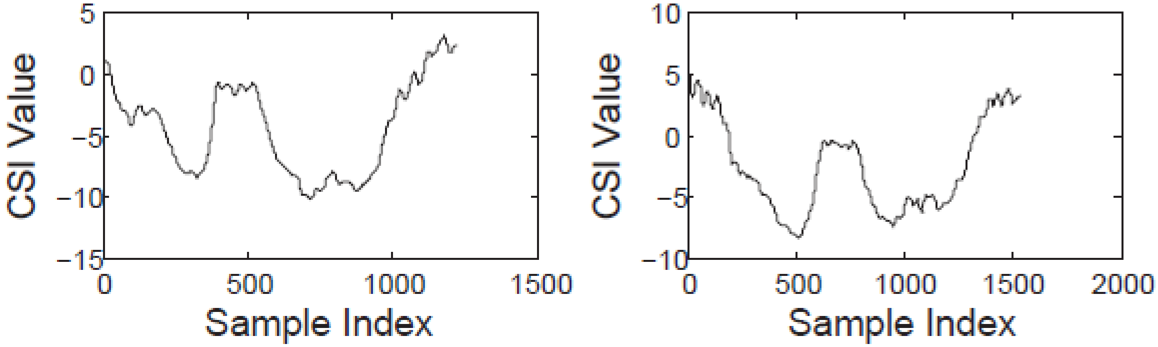
Extraction

Password Inference Module

- Keystroke Recognition



(a) Two samples of keystroke waveforms number 2

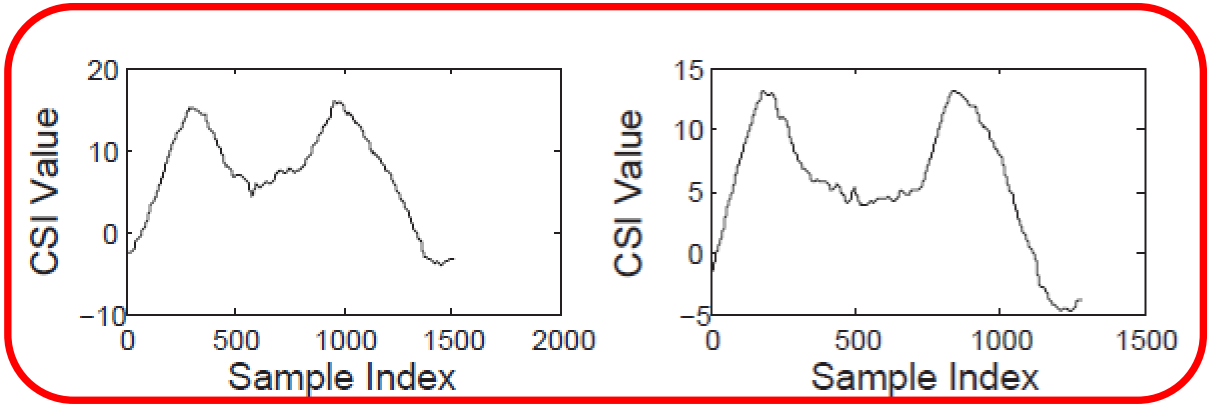


(b) Two samples of keystroke waveforms number 4

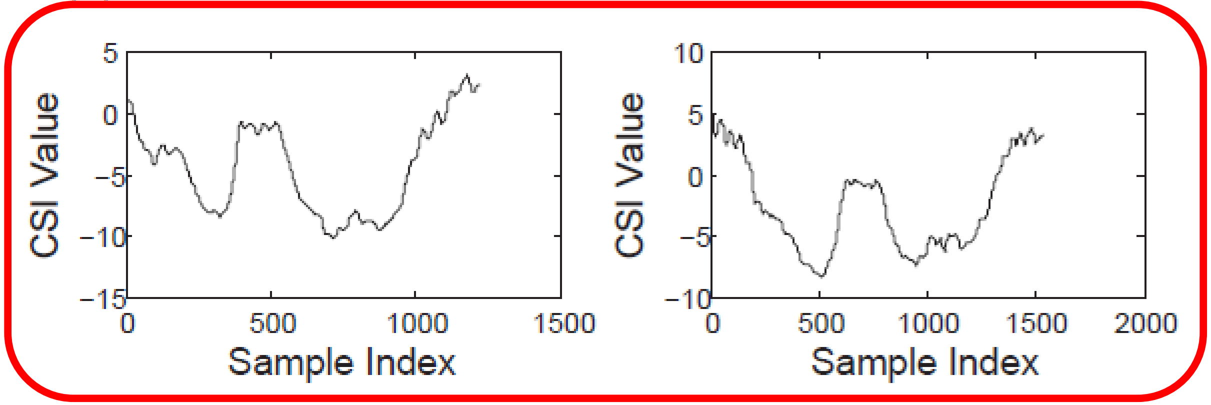
- Dynamic Time Warping
- Classifier Training
- Recognition

Password Inference Module

Keystroke Recognition



(a) Two samples of keystroke waveforms number 2



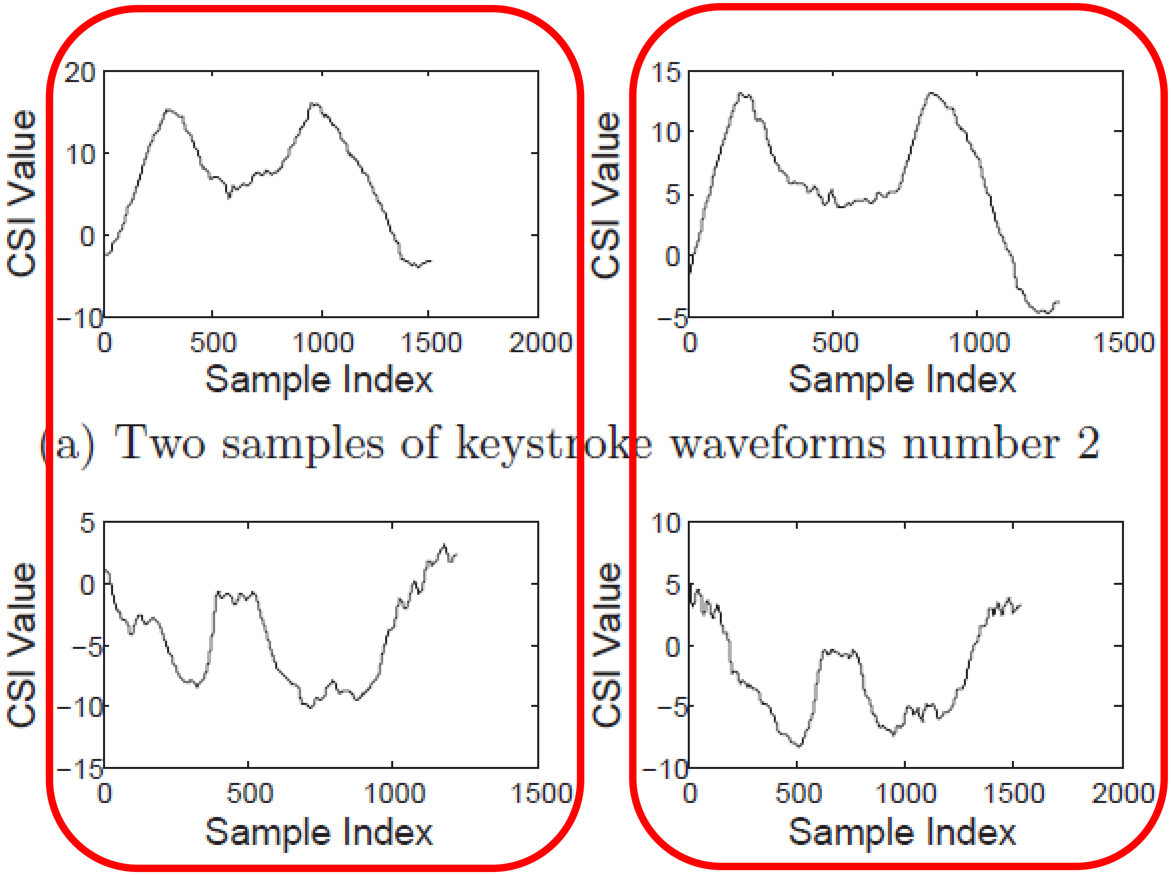
(b) Two samples of keystroke waveforms number 4

Same Number

DTW Distance

Password Inference Module

- Keystroke Recognition



(a) Two samples of keystroke waveforms number 2

(b) Two samples of keystroke waveforms number 4

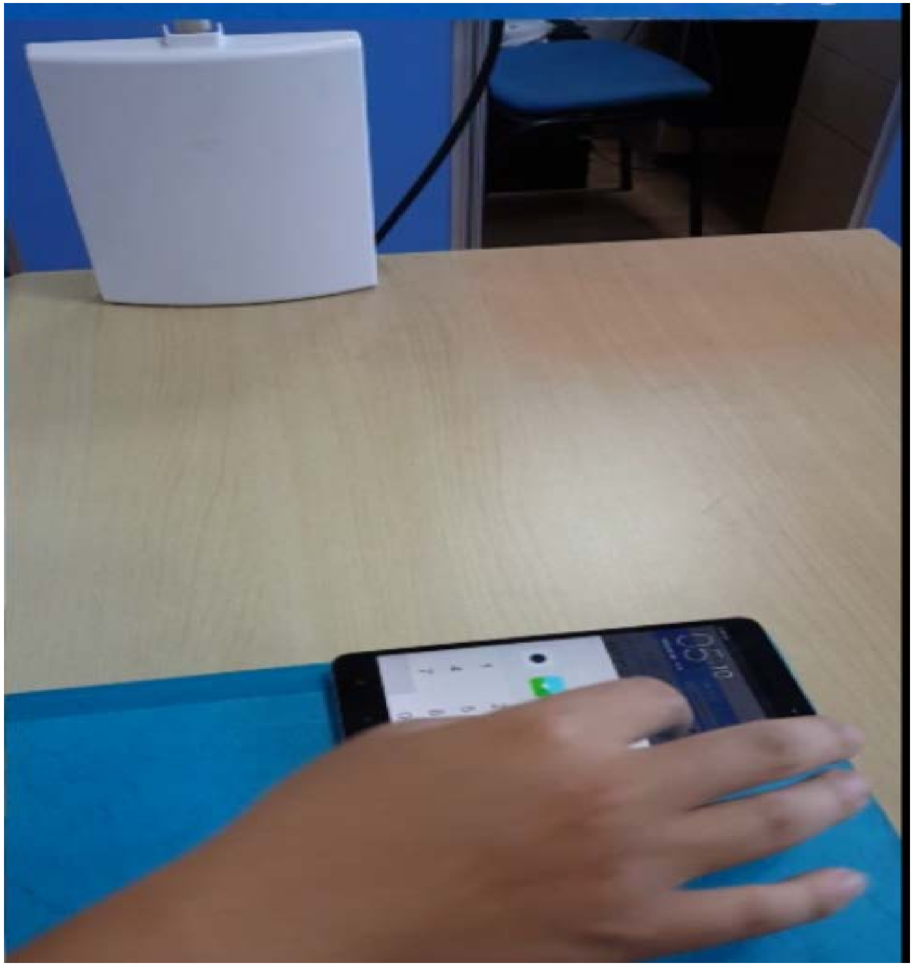
Different Number

DTW Distance ↑

OUTLINE

- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

Classification between Different Numbers



- 10 Volunteers

  3 Types of Phone

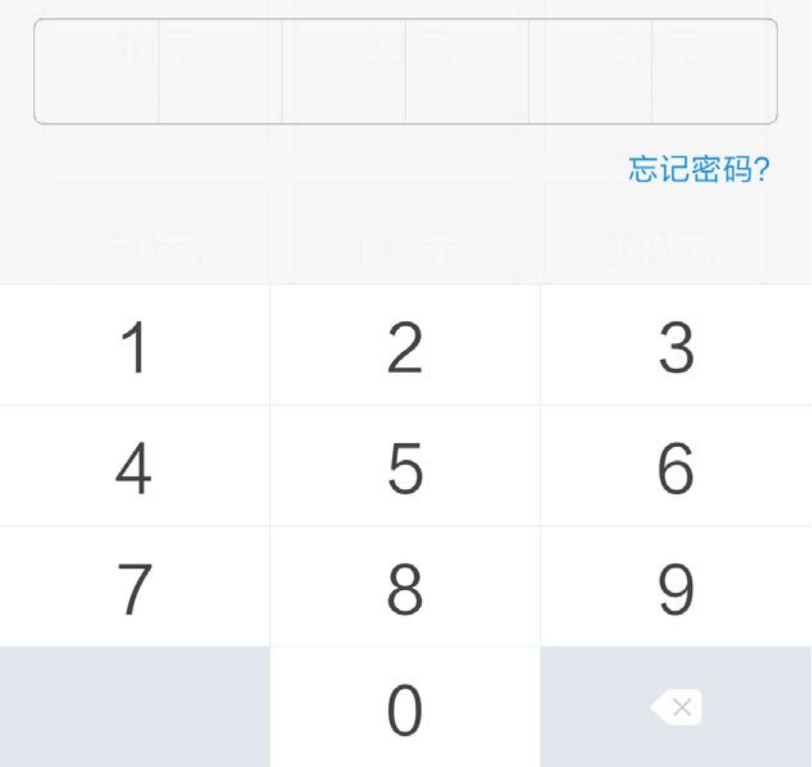
- Each Volunteer:

  Press 10 Loops

- Each Loop:

  from 1-2-3-…-0

Classification between Different Numbers



- 10 Volunteers

  3 Types of Phone

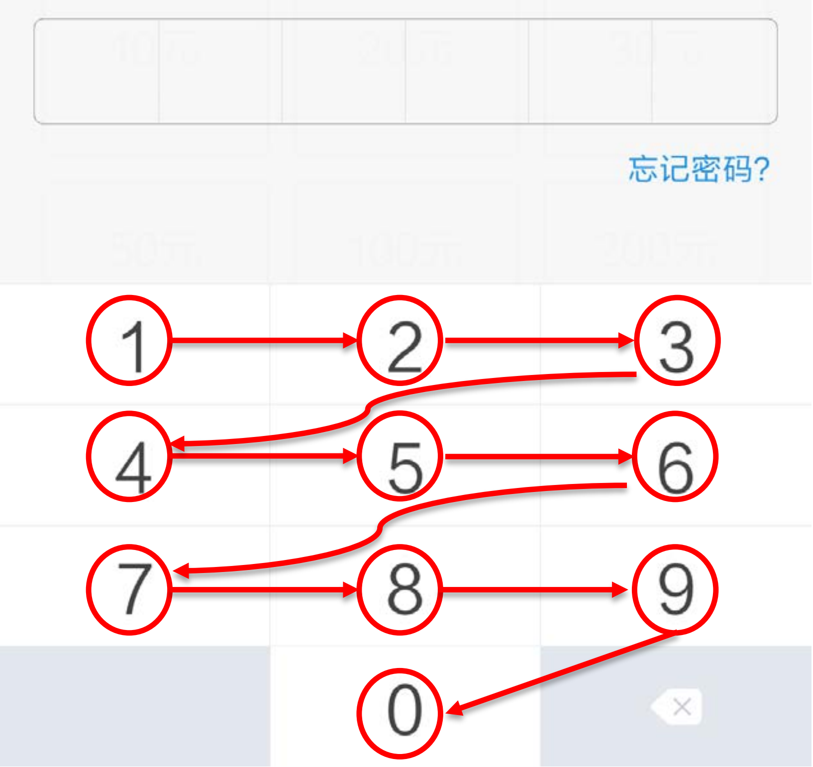
- Each Volunteer:

  Press 10 Loops

- Each Loop:

  from 1-2-3-….-0

Classification between Different Numbers



- 10 Volunteers

  3 Types of Phone

- Each Volunteer:

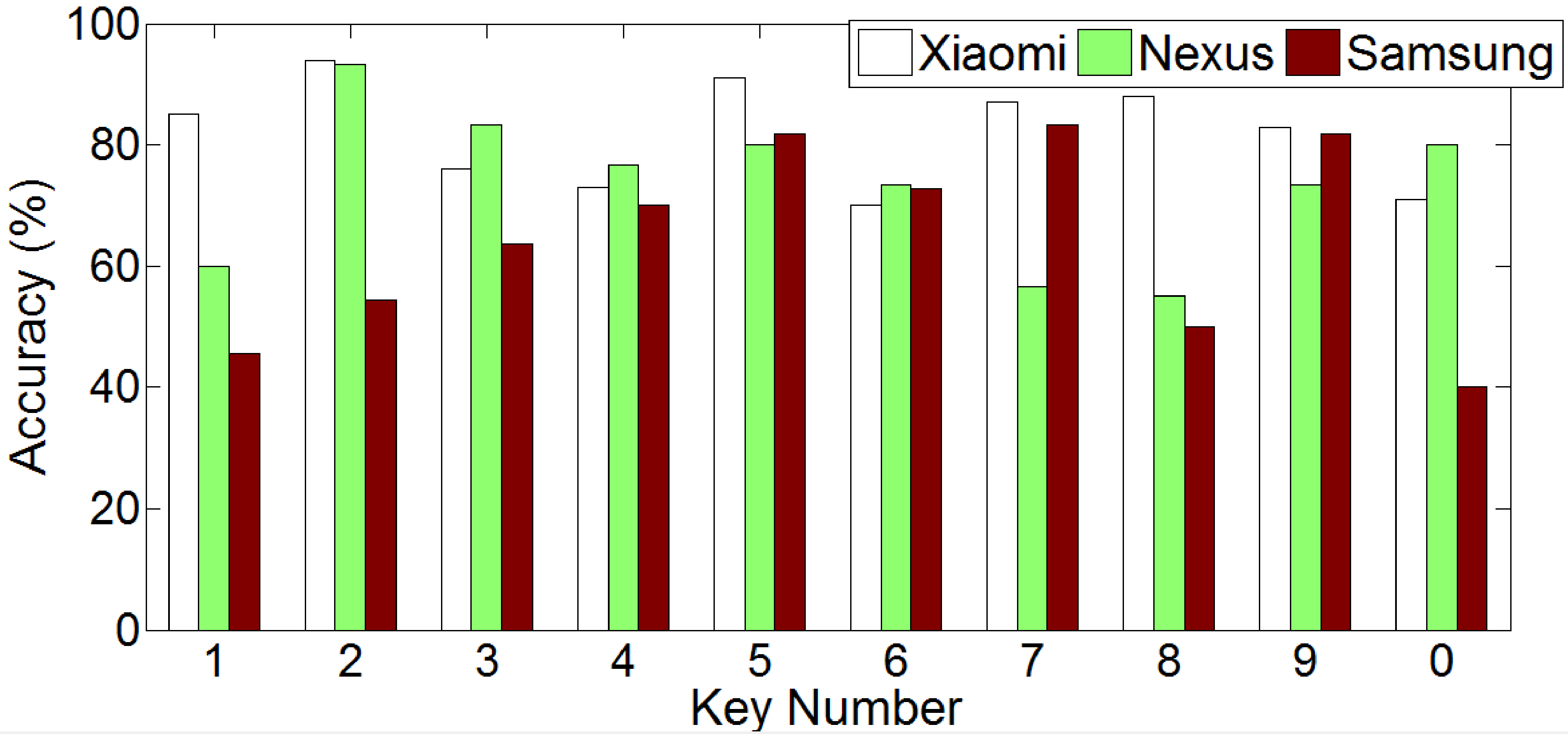
  Press 10 Loops

- Each Loop:

  from 1-2-3-….-0

Classification between Different Numbers

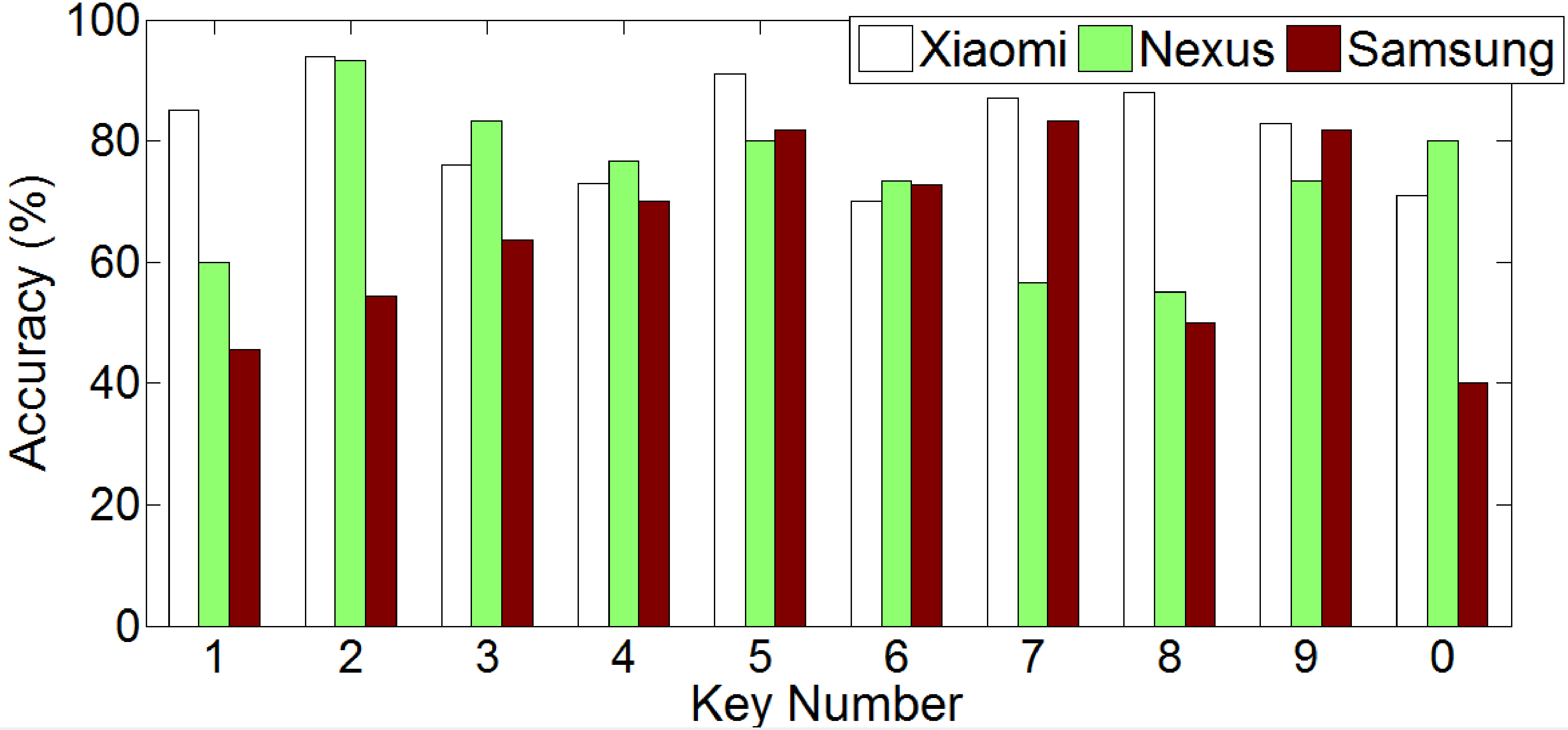
- Classification Results:



Cross validation accuracy. Each times, 1 loop for testing and 9 loops for training.

Classification between Different Numbers

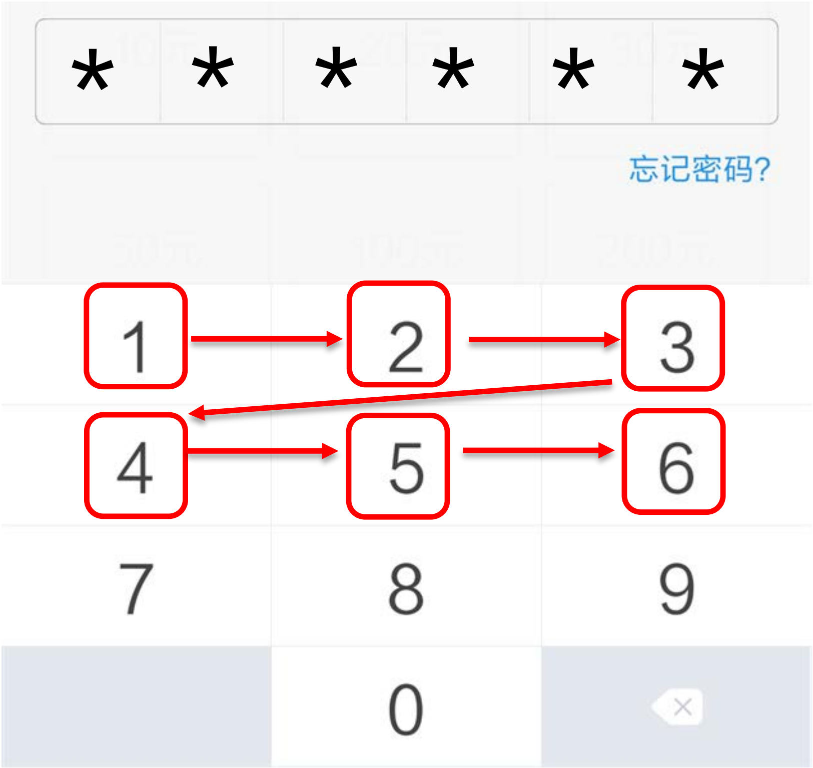
- Classification Results:



- 82% in Xiaomi, 73% in Nexus and 64% in Samsung

Infer 6-digit password

6-digit password is a fixed password format for Alipay, Wechat pay and many other online banks.



Use Password Candidates

Possible candidates for “123456”

125484

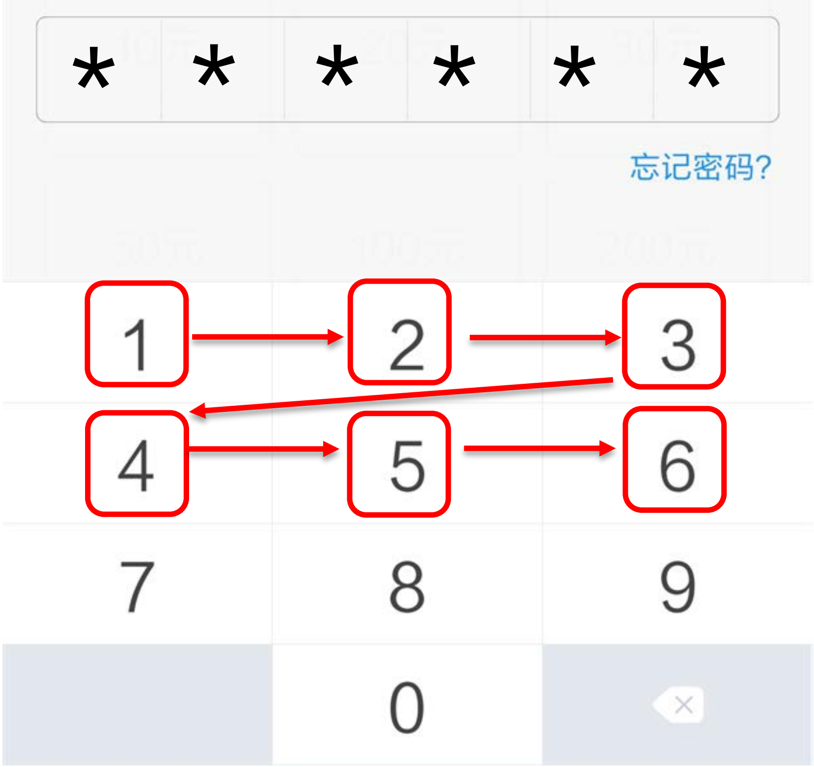
215487

123456

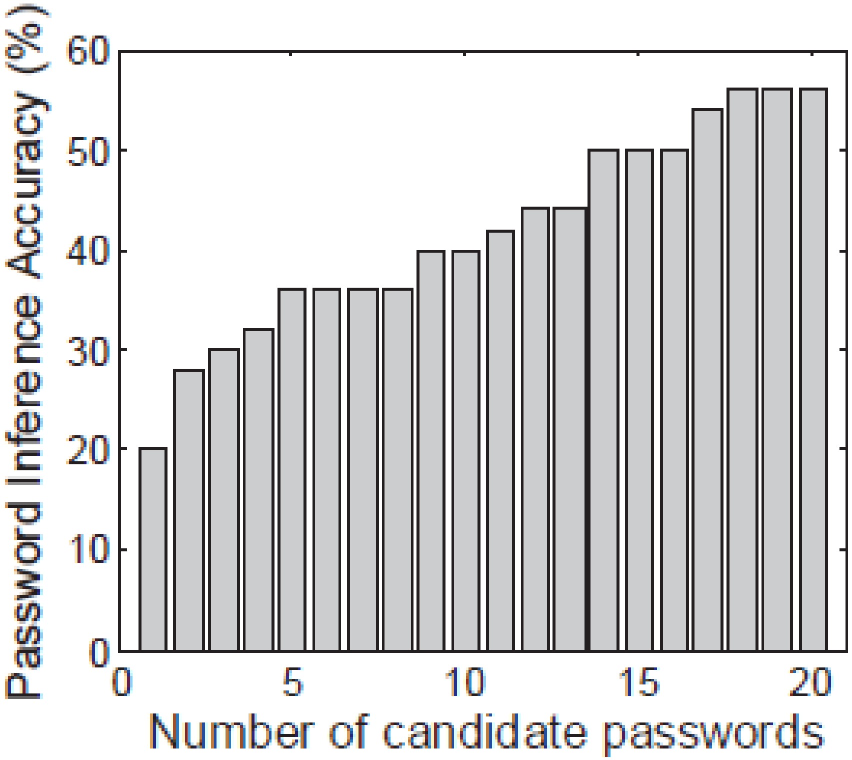
……

Infer 6-digit password

6-digit password is a fixed password format for Alipay, Wechat pay and many other online banks.



Use Password Candidates

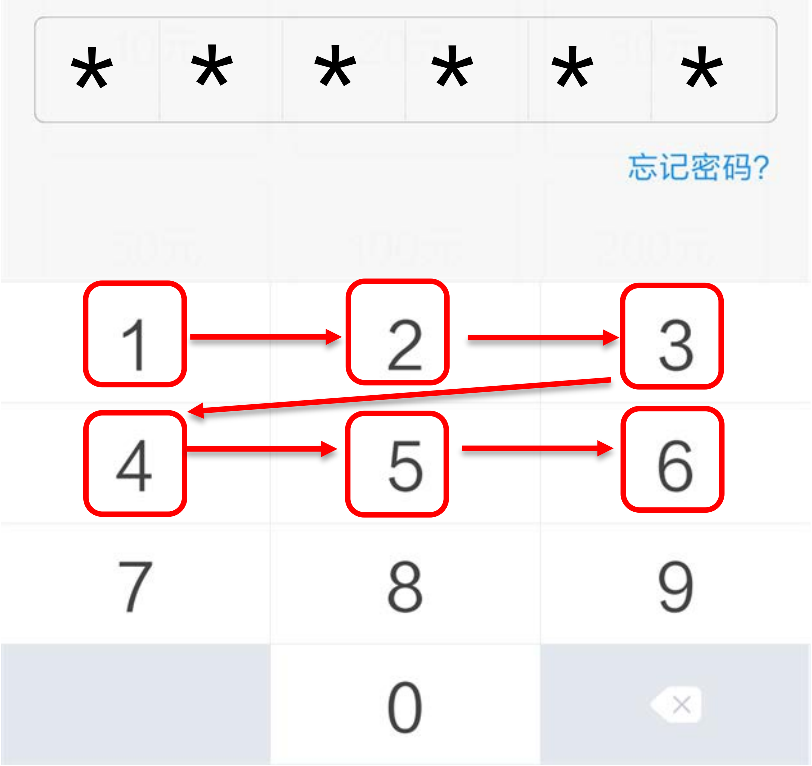


3 Loops for training

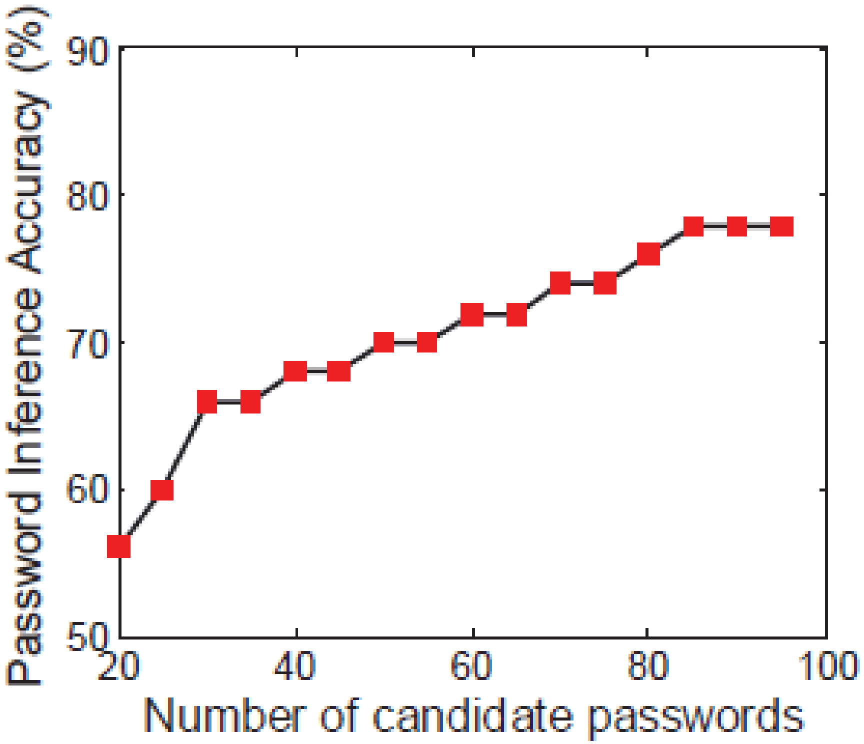
200 passwords from ten volunteers

Infer 6-digit password

6-digit password is a fixed password format for Alipay, Wechat pay and many other online banks.



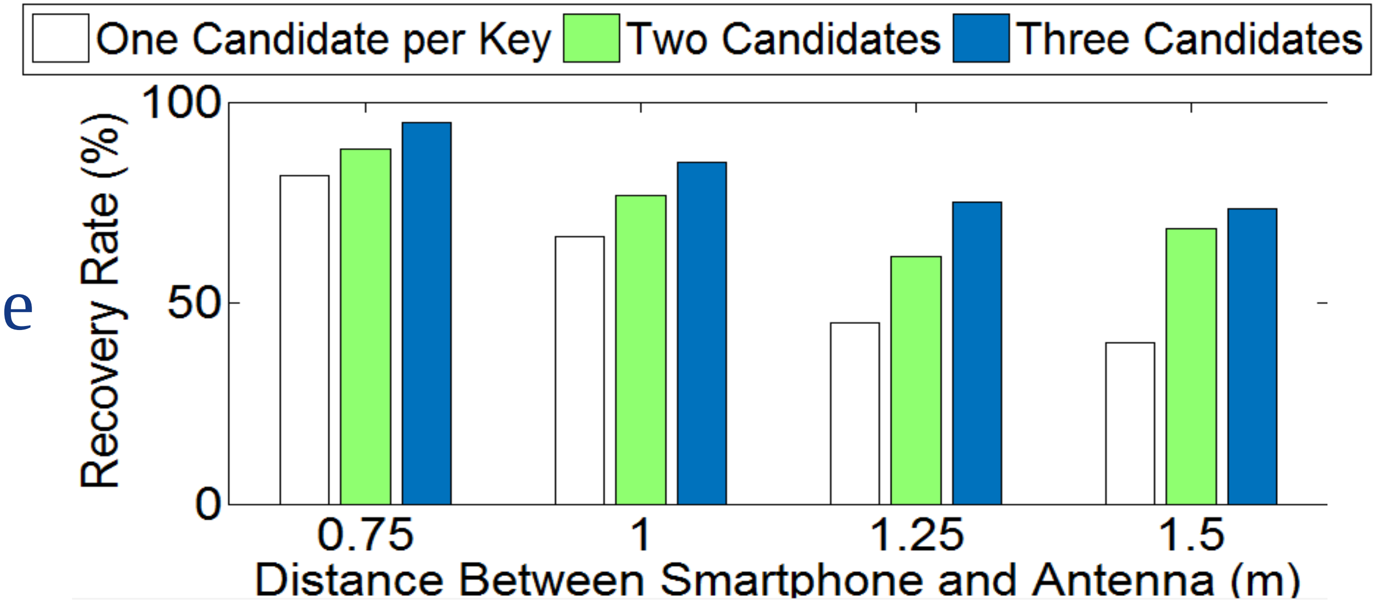
Use Password Candidates



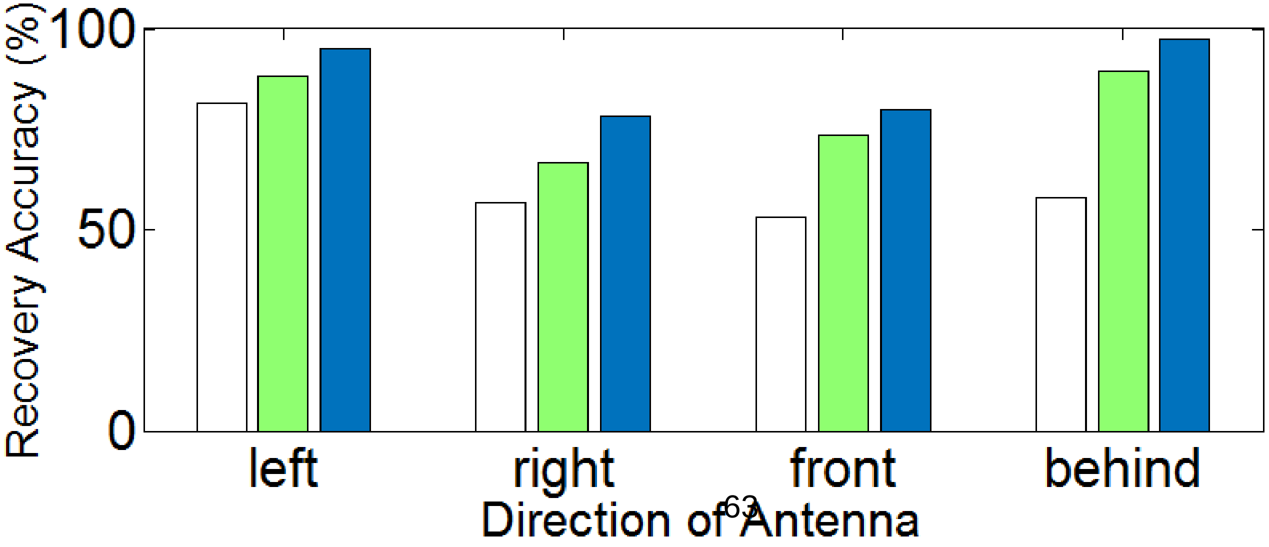
3 Loops for training
200 passwords from ten volunteers

Influence factors

Evaluation on Different Distance



Evaluation on Different Direction

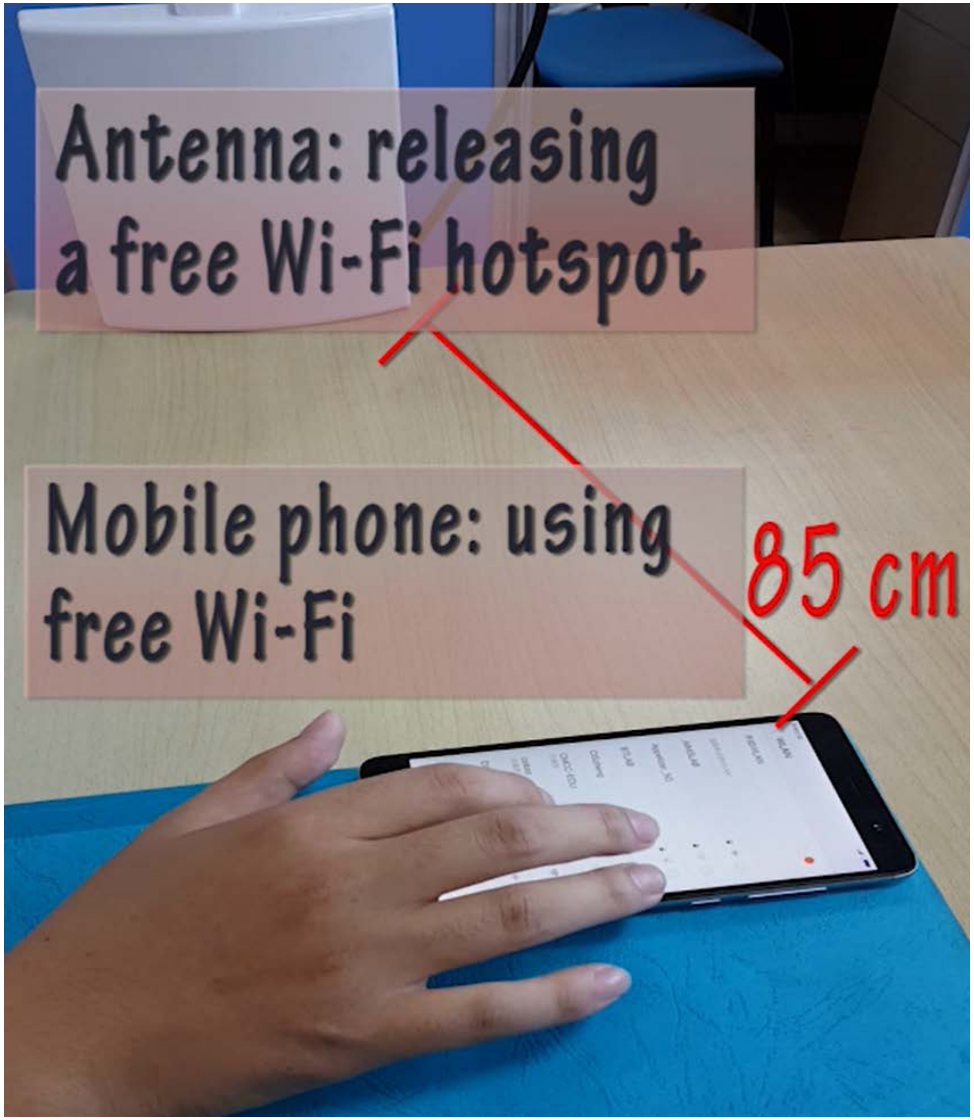


OUTLINE

- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

Case Study

- Simulate Real-world Scenario
- Combine Four Technical Modules
- Click **Demo** to See Details


Antenna: releasing a free Wi-Fi hotspot
Mobile phone: using free Wi-Fi
85 cm

Case Study

- Simulate Real-world Scenario
- Combine Four Technical Modules
- Click **Demo** to See Details
- Case Study Results

Carry out case study 10 times:

Candidates Number	**Successfully Inference**
5	2
10	4
50	7
100	9

OUTLINE

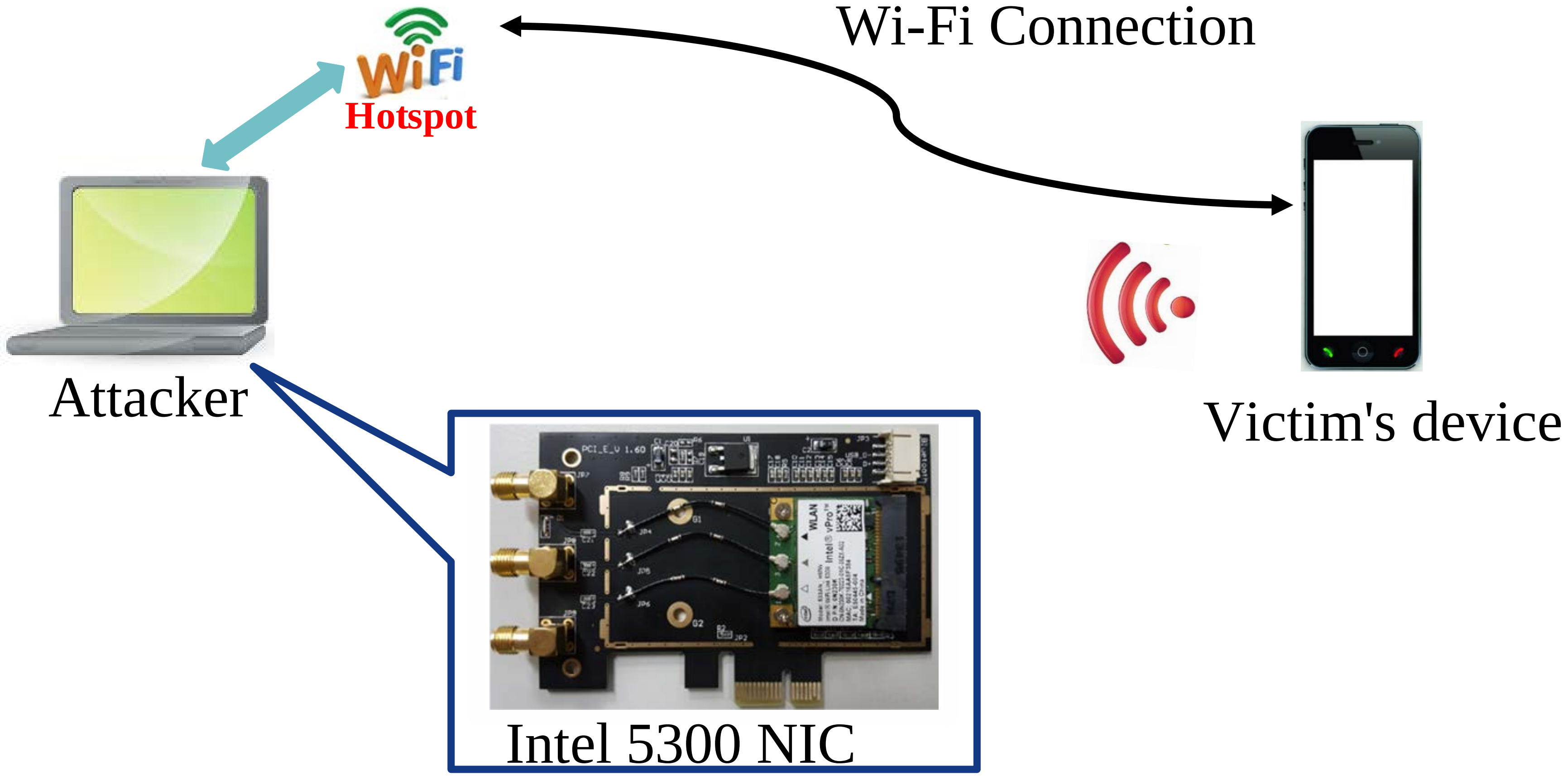
- Motivation
- Attack Scenario
- System Design
- Evaluation
- Case Study
- Conclusion

Limitations

- Hardware Limitations
- Fixed Typing Gesture
- User Specific Training

Limitations

- Hardware Limitations



Wi-Fi Connection

Hotspot

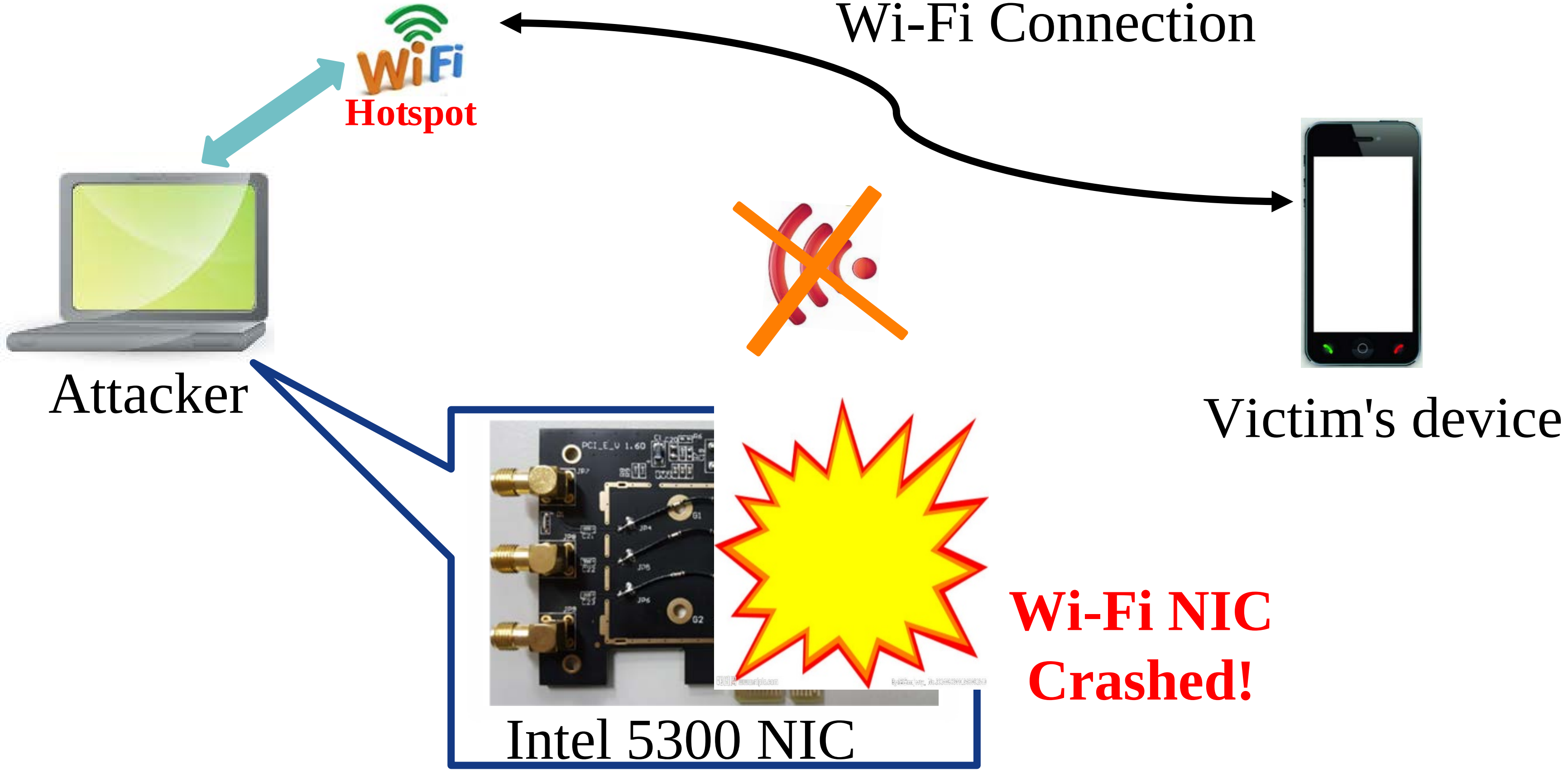
Attacker

Victim's device

Intel 5300 NIC

Limitations

- Hardware Limitations



Hotspot

Wi-Fi Connection

Attacker

Victim's device

Intel 5300 NIC

Wi-Fi NIC Crashed!

Limitations

- Hardware Limitations
- Fixed Typing Gesture

  Too quick type

  Strange hand motion

  Disturbance nearby

Limitation

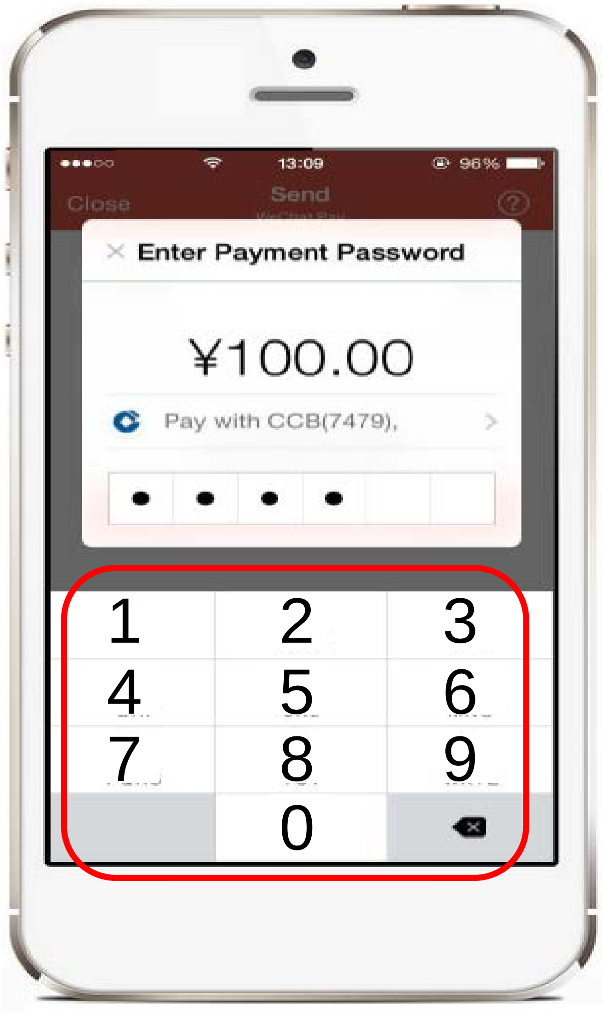
- Hardware Limitations
- Fixed Typing Gesture
- User Specific Training

  Text Captchas

  Plain content analysis

Countermeasure

- Random Layouts of Keyboard



●●●○○ 13:09 96%

Close Send ?

× Enter Payment Password

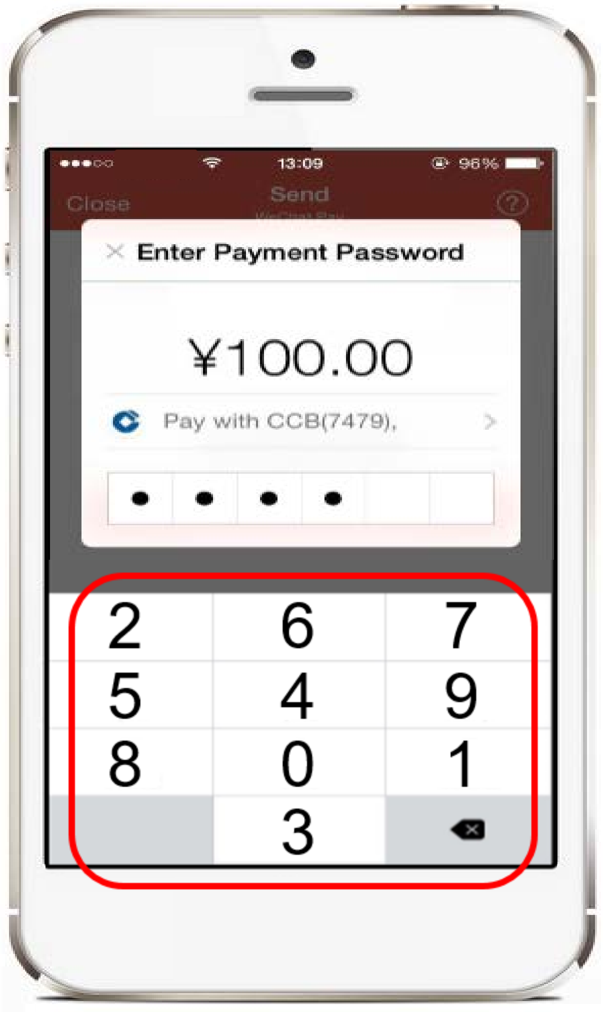
¥100.00

Pay with CCB(7479),

● ● ● ●

1	2	3
4	5	6
7	8	9
	0	⌫

After typing



●●●○○ 13:09 96%

Close Send ?

× Enter Payment Password

¥100.00

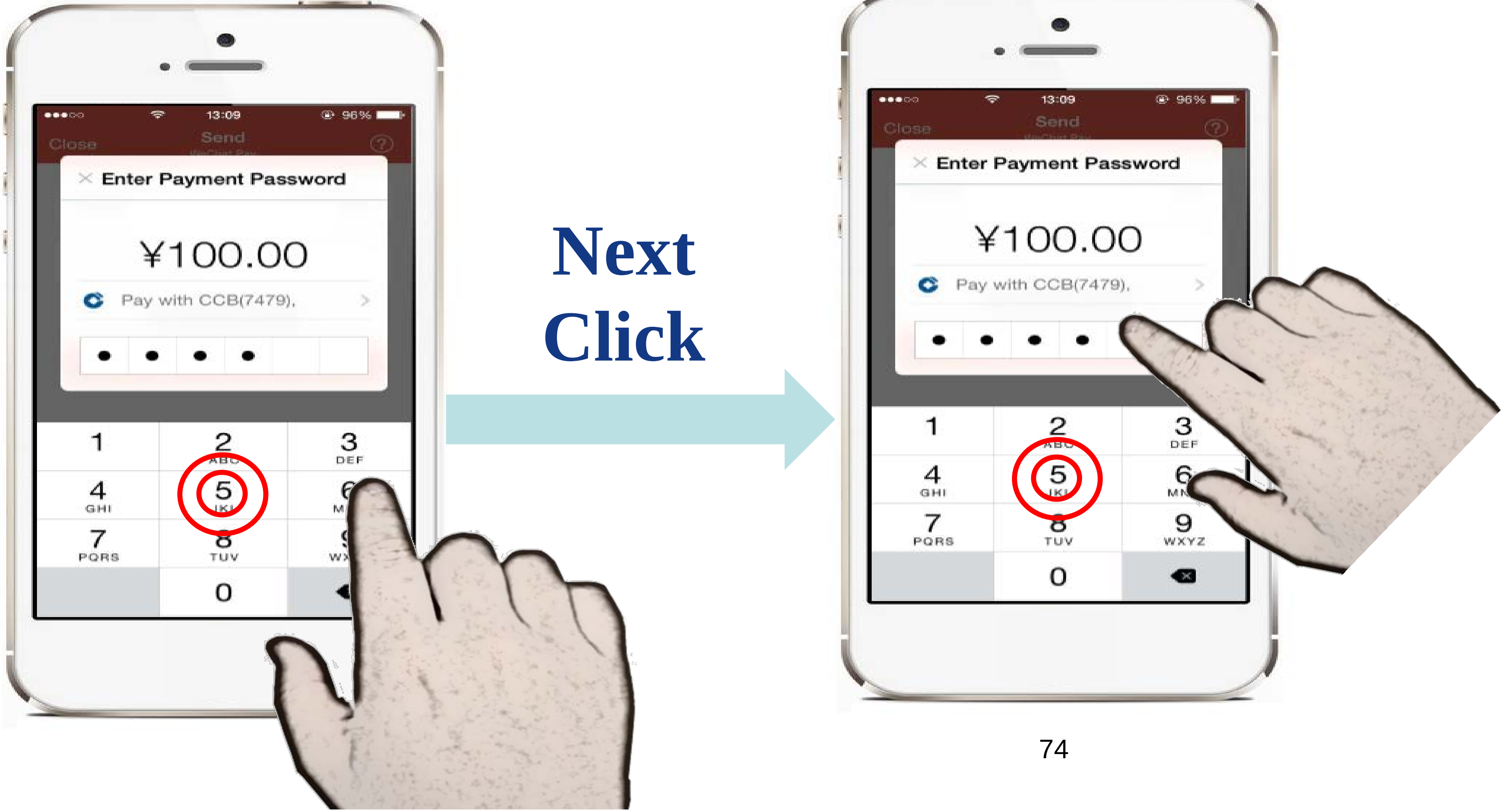
Pay with CCB(7479),

● ● ● ●

2	6	7
5	4	9
8	0	1
	3	⌫

Countermeasure

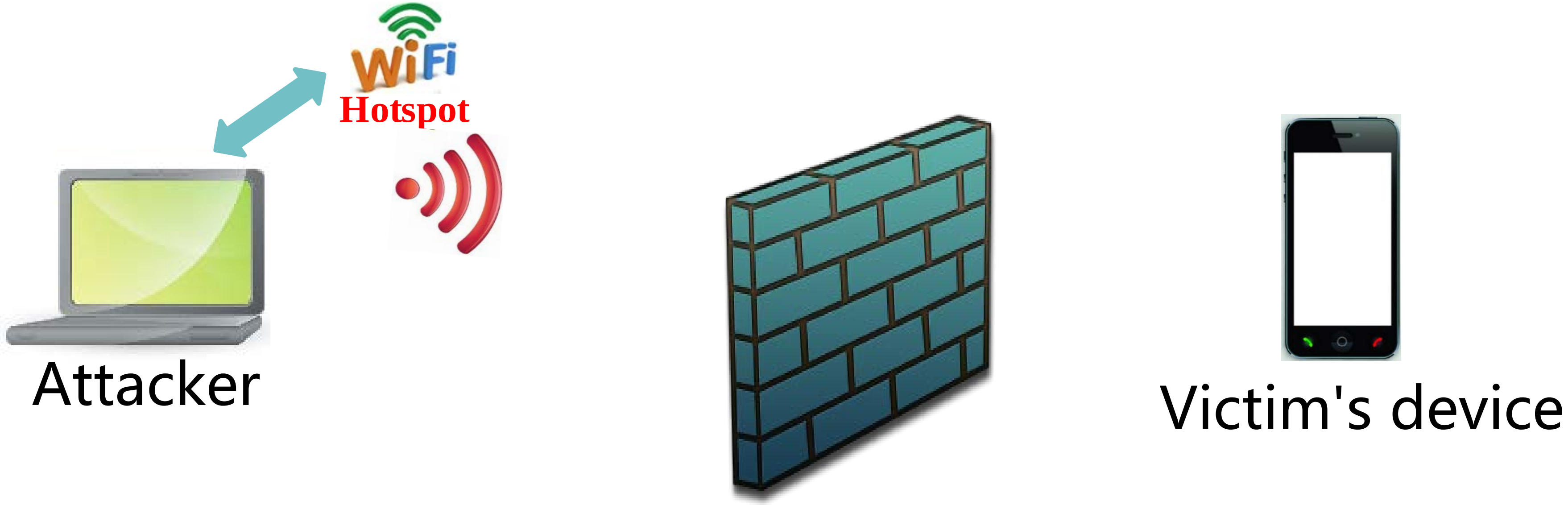
- Random Layouts of Keyboard
- Change Typing Gesture



Next Click

Countermeasure

- Random Layouts of Keyboard
- Change Typing Gesture
- Preventing the collection of CSI



Conclusion and Future Work

- We present WindTalker, a novel attack that uses physical layer information to attack applications in the upper layers (Encryption may not work).

- It is expected to have a broad potential application for password inference in mobile devices (encrypted traffic analysis + CSI analysis should be cool).

- Major issue is the CSI collection module is not reliable: using advanced tools to enhance it.

Thank you!

Haojin Zhu
Zhu-hj@cs.sjtu.edu.cn