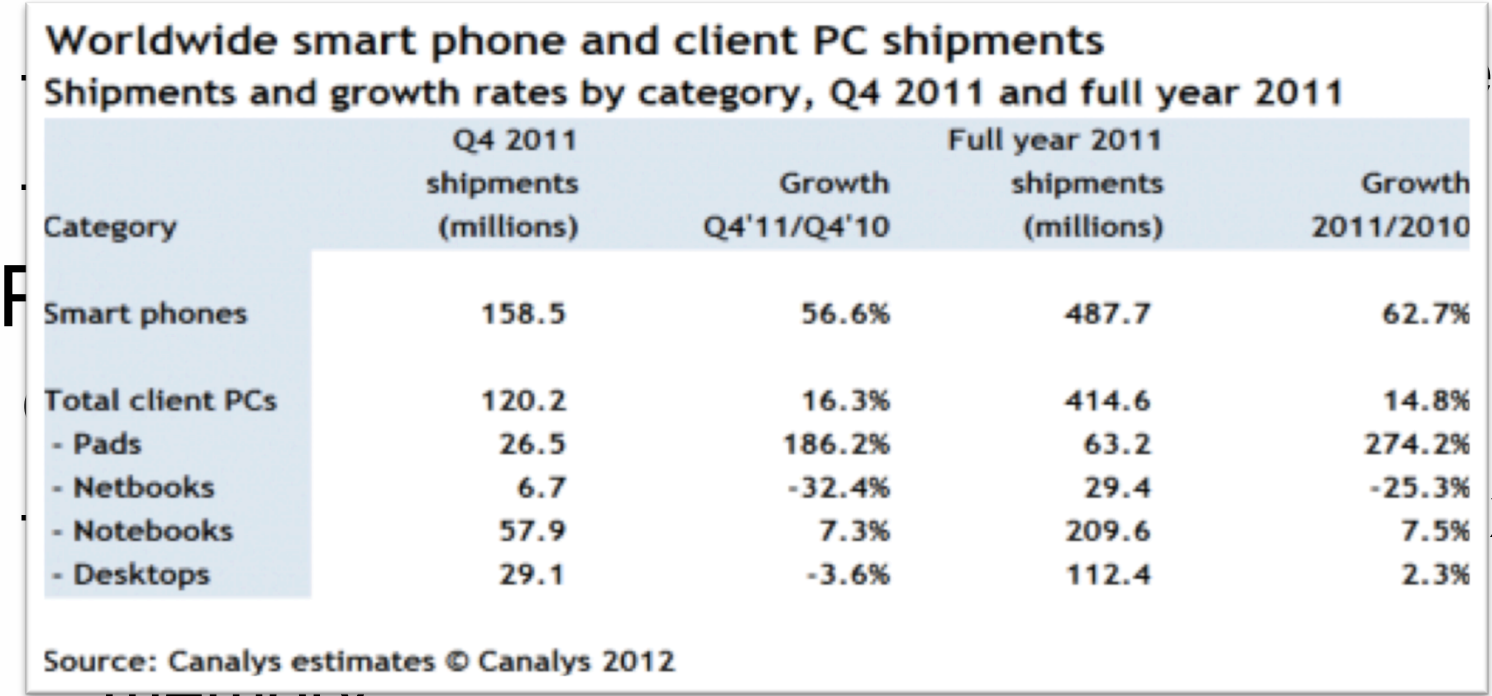
Toward a Trustworthy Android Ecosystem

Yan Chen
Lab of Internet and Security Technology (LIST)
Northwestern University, USA

Smartphone Security

- Ubiquity - Smartphones and mobile devices



Worldwide smart phone and client PC shipments

Shipments and growth rates by category, Q4 2011 and full year 2011

Category	Q4 2011 shipments (millions)	Growth Q4'11/Q4'10	Full year 2011 shipments (millions)	Growth 2011/2010
Smart phones	158.5	56.6%	487.7	62.7%
Total client PCs	120.2	16.3%	414.6	14.8%
- Pads	26.5	186.2%	63.2	274.2%
- Netbooks	6.7	-32.4%	29.4	-25.3%
- Notebooks	57.9	7.3%	209.6	7.5%
- Desktops	29.1	-3.6%	112.4	2.3%

Source: Canalys estimates © Canalys 2012

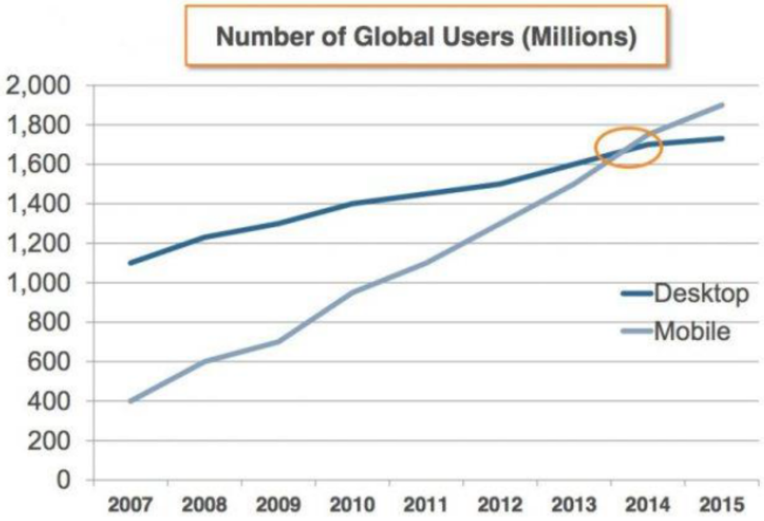
es

- F

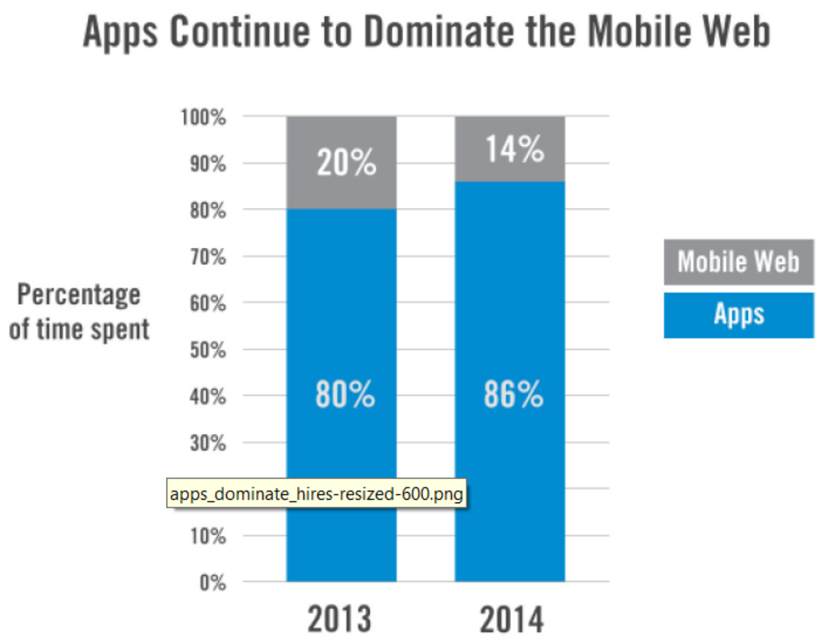
xel

memory

Mobile Devices (apps) Dominate

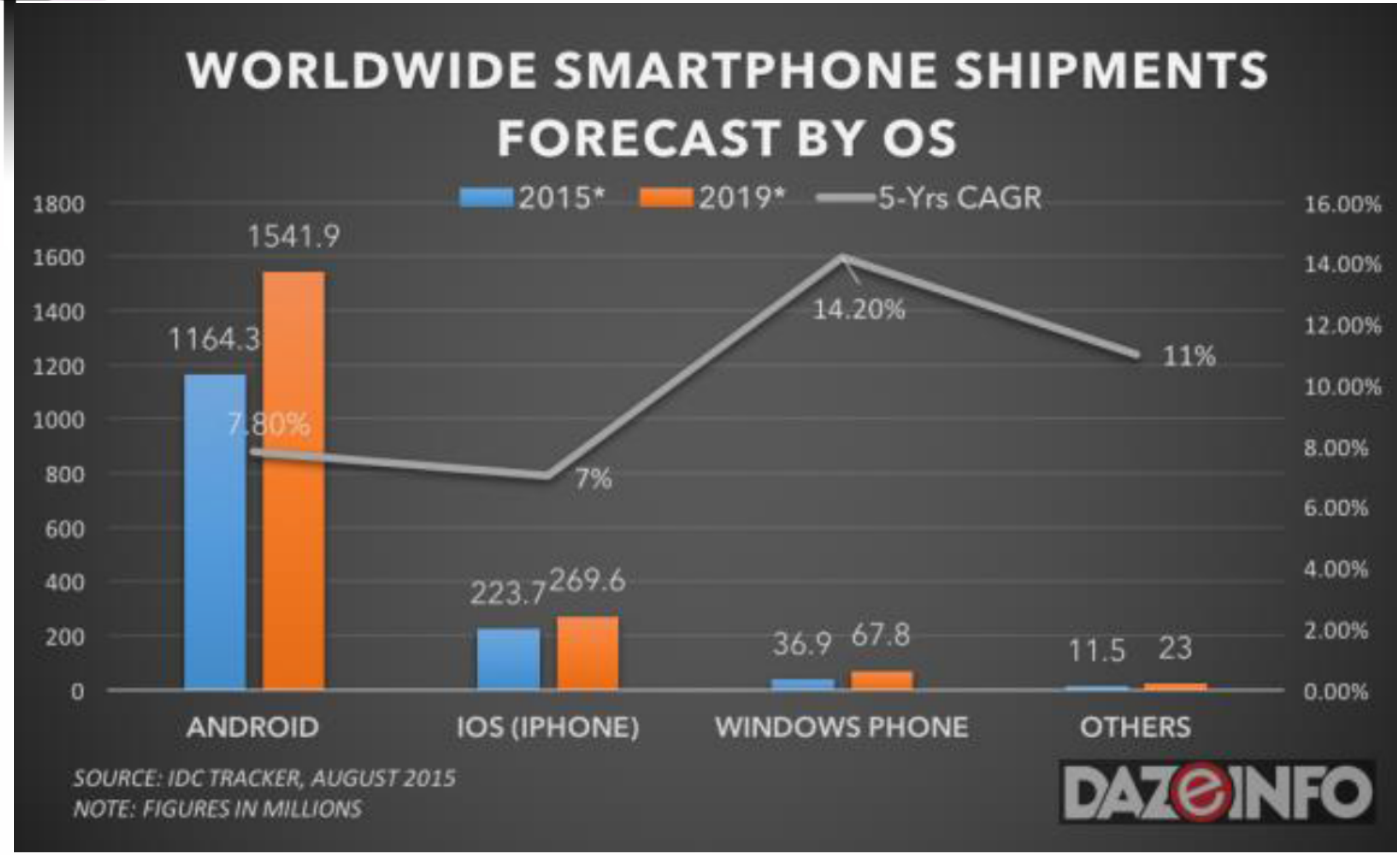


comScore. © comScore, Inc. Proprietary and Confidential. 24 Source: Morgan Stanley Research



FLURRY Source: Flurry Analytics

Android is Leading the Pack



SOURCE: IDC TRACKER, AUGUST 2015
NOTE: FIGURES IN MILLIONS

Android Ecosystem

Carriers

at&t

verizon

Sprint

T-Mobile

Vendors

SAMSUNG

LG

Application Stores

Google play

Google

Devices and OS

Applications

KASPERSKY ANTI-VIRUS

AVG

McAfee

ca INTERNET SECURITY SUITE 2010

bitdefender

Microsoft Security Essentials

F-Secure

avast! be free

AVIRA

eset NOD32 antivirus

Security Vendors

Users

Developers

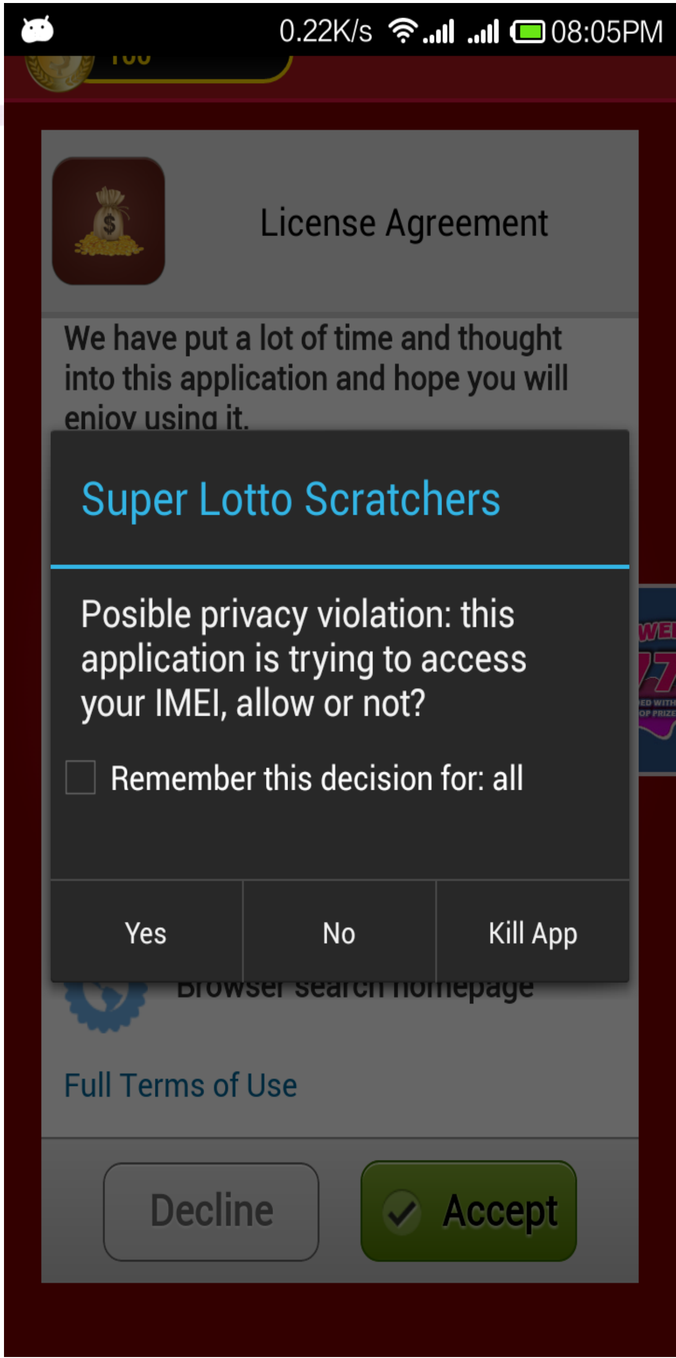
Android Threats

flickr.com/photos/panda_security_france/

- Malware and vulnerabilities
  - The numbers are increasing consistently
  - Anti-malware ineffective at catching zero-day and polymorphic malware
- Information Leakage
  - Users have no way to know when and what info is being leaked out of their device to whom
  - Even legitimate apps leak private info though the user may not be aware
- Fraud activities (esp. for mobile payment)

Privacy Leakage

- Android permissions are insufficient
  - User still does not know if some private information will be leaked
- Information leakage is more dangerous than information access
  - Example 1: popular apps (e.g., Angry Birds) leak location info with its developer, advertisers and analytics services
    - Even doesn’t need it for its functionality!
  - Example 2: malware apps may steal private data
    - A camera app trojan send video recordings out of the phone



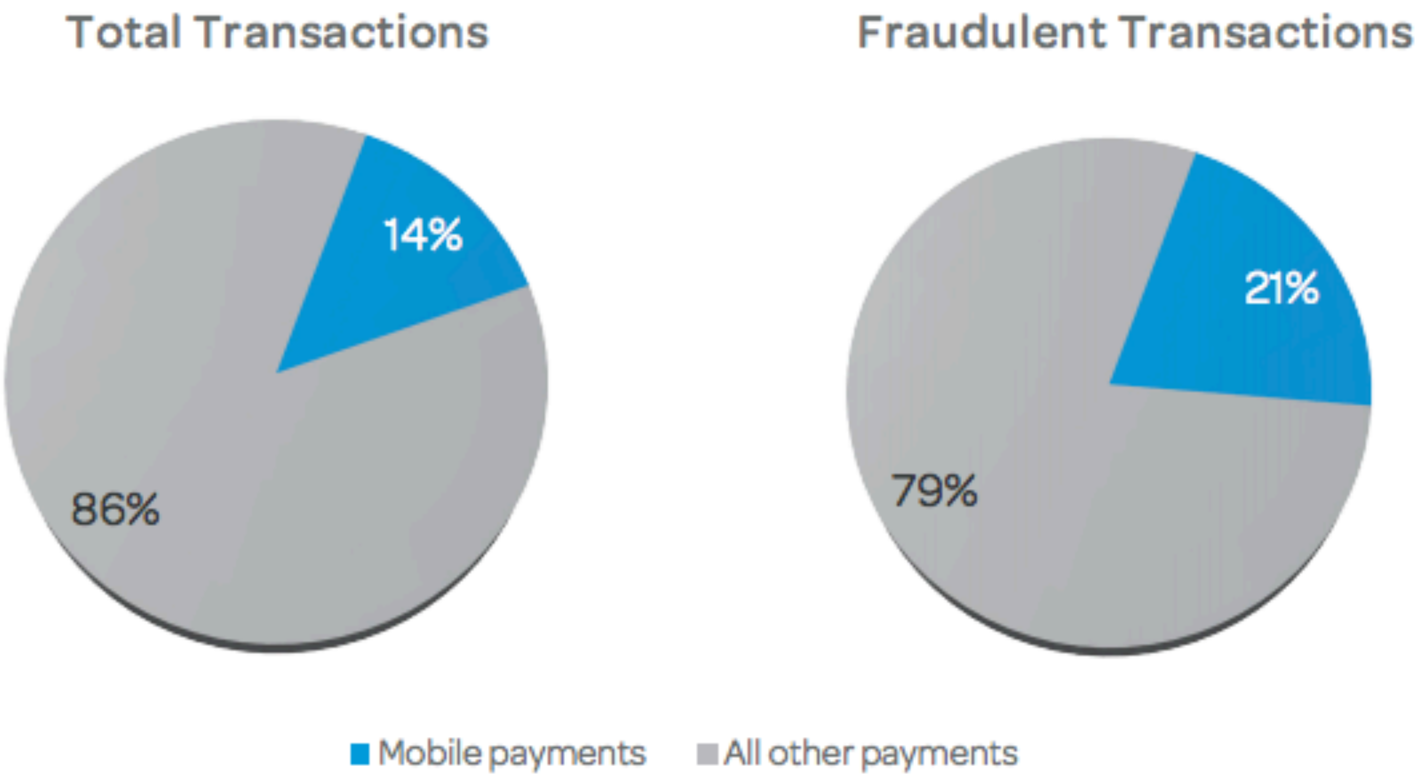
We have put a lot of time and thought into this application and hope you will enjoy using it.

Super Lotto Scratchers

Posible privacy violation: this application is trying to access your IMEI, allow or not?

☐ Remember this decision for: all

Fraudulent Mobile Transactions



123,255 eCommerce customers with >$0 fraud in March 2014, by LexisNexis

New Challenges & Opportunities

- Centralized control
  - – Vet applications before they enter store
  - – Carriers may have more complete pictures of users and traffic
- Apps are much easier to analyze statically
  - – Use of Dalvik bytecode instead of x86
- Constrained environment
  - – CPU, memory, battery
  - – User perception

Problems and Our Solutions

- Issues for existing mobile anti-virus systems
  - Easy to evade [DroidChamelon]
  - Unable to detect native malware [DroidNative]
  - Unable to detect malware in ads or dynamically loaded content [AdShield]

- Privacy leakage detection and prevention
  - How to find questionable sensitive permissions [AutoCog]
  - Real time tracking & preventing privacy leakage on phone
    - Consumer [PrivacyShield]
    - Enterprise Mobility Management (EMM) [AppShield]

- Fraud detection mostly with app-level risk management [DroidCog]
  - Duplicate detection
  - Privacy infringement

Systems Developed

- AppsPlayground [ACM CODASPY'13]
  - Automatic, large-scale dynamic analysis of Android apps
  - System released with hundreds of download
- DroidChameleon [ACM ASIACCS'13, IEEE Transaction on Information Forensics and Security 14]
  - Evaluation of latest Android anti-malware tools
  - All can be evaded with transformed malware
  - System released upon wide interest from media and industry

Impact of DroidChamelon

heise online

McCormick Northwestern Engineering

THE WALL STREET JOURNAL.

Security darkREADING Protect The Business Enable Access

Slashdot

InformationWeek THE BUSINESS VALUE OF TECHNOLOGY

SECURITYWEEK INTERNET AND ENTERPRISE SECURITY NEWS, INSIGHTS & ANALYSIS

PHYS.ORG

NBCNEWS.com

ScienceDaily® Your source for the latest research news

TECHNEWSDAILY

ISS Source

EFYTimes.com

VirusFreePhone.com Protecting your phone from malware and viruses

ACM TechNews

HELP NET SECURITY

Interest from vendors

Lookout

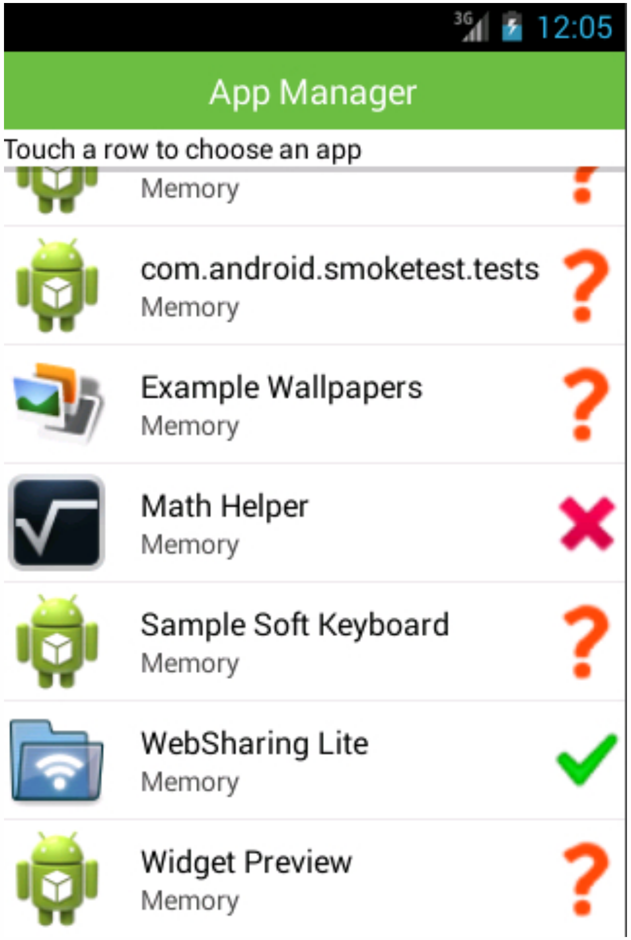
McAfee An Intel Company

N

AVG

System Developed: AutoCog

Check whether sensitive permissions requested by apps are consistent with its natural-language description



Touch a row to choose an app

Memory ?

com.android.smoketest.tests
Memory ?

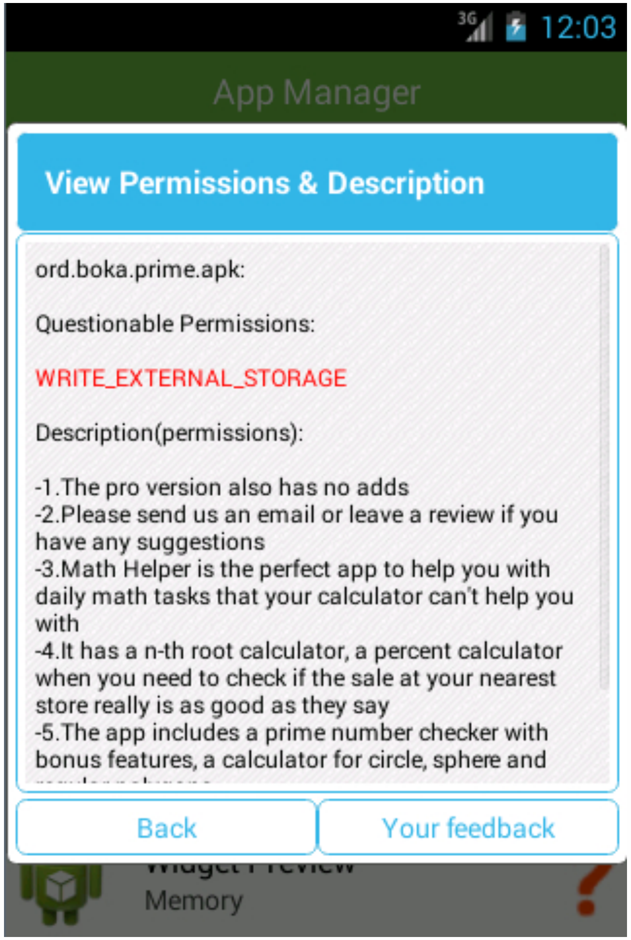
Example Wallpapers
Memory ?

Math Helper
Memory ✗

Sample Soft Keyboard
Memory ?

WebSharing Lite
Memory ✓

Widget Preview
Memory ?



ord.boka.prime.apk:

Questionable Permissions:

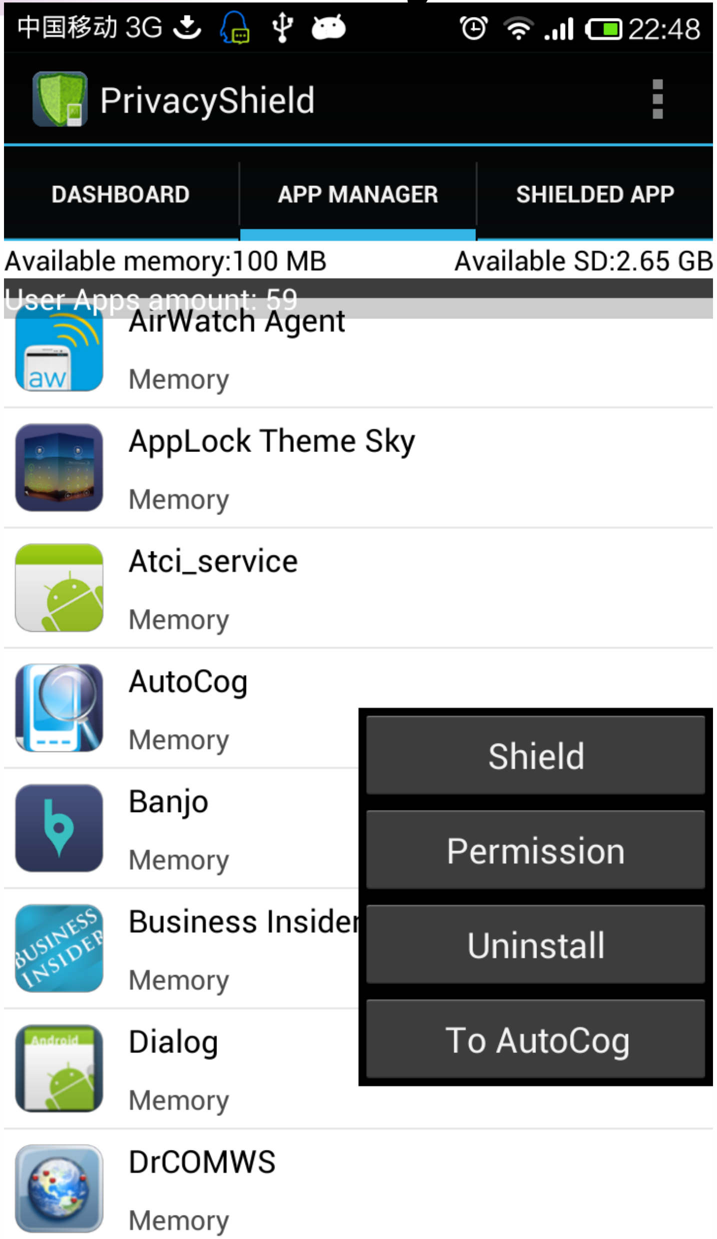
WRITE_EXTERNAL_STORAGE

Description(permissions):

-1.The pro version also has no adds
-2.Please send us an email or leave a review if you have any suggestions
-3.Math Helper is the perfect app to help you with daily math tasks that your calculator can't help you with
-4.It has a n-th root calculator, a percent calculator when you need to check if the sale at your nearest store really is as good as they say
-5.The app includes a prime number checker with bonus features, a calculator for circle, sphere and

Systems Developed: PrivacyShield

- Real-time information -flow tracking for privacy leakage detection
- App instrumentation, with zero platform modification
- App released in Google play and Baidu stores



Available memory:100 MB
Available SD:2.65 GB
User Apps amount: 59

AirWatch Agent
Memory

AppLock Theme Sky
Memory

Atci_service
Memory

AutoCog
Memory

Banjo
Memory

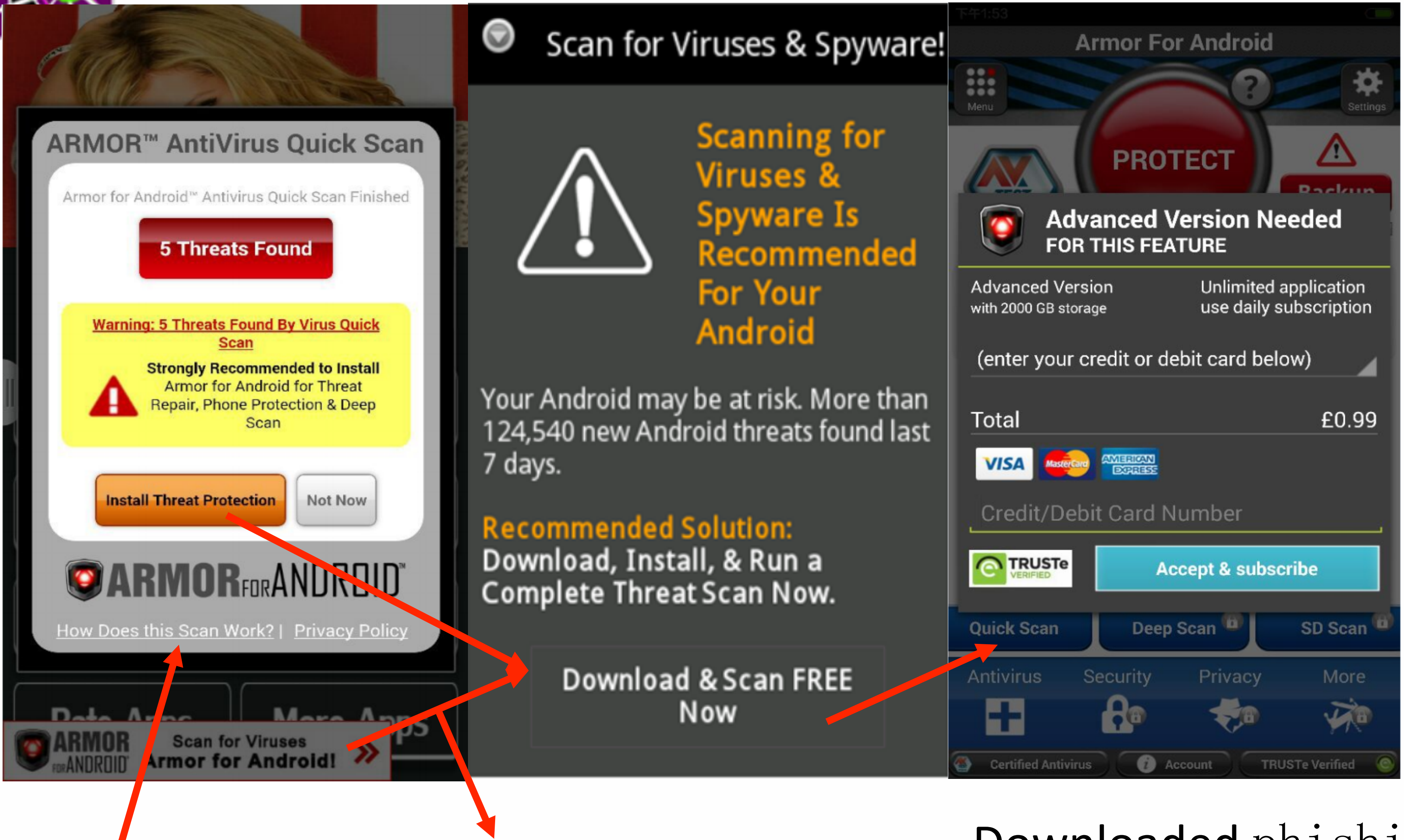
Business Insider
Memory

Dialog
Memory

DrCOMWS
Memory

ARE THESE ADS SAFE: DETECTING HIDDEN ATTACKS THROUGH MOBILE APP-WEB INTERFACES

Consider This…



ARMOR™ AntiVirus Quick Scan

Armor for Android™ Antivirus Quick Scan Finished

5 Threats Found

Warning: 5 Threats Found By Virus Quick Scan

Strongly Recommended to Install Armor for Android for Threat Repair, Phone Protection & Deep Scan

Install Threat Protection | Not Now

ARMOR FOR ANDROID™

How Does this Scan Work? | Privacy Policy

ARMOR FOR ANDROID Scan for Viruses Armor for Android! »

Scan for Viruses & Spyware!

Scanning for Viruses & Spyware Is Recommended For Your Android

Your Android may be at risk. More than 124,540 new Android threats found last 7 days.

Recommended Solution: Download, Install, & Run a Complete Threat Scan Now.

Download & Scan FREE Now

Armor For Android

Menu | PROTECT | Settings

Advanced Version Needed FOR THIS FEATURE

Advanced Version with 2000 GB storage | Unlimited application use daily subscription

(enter your credit or debit card below)

Total £0.99

VISA MasterCard AMERICAN EXPRESS

Credit/Debit Card Number

TRUSTe VERIFIED | Accept & subscribe

Quick Scan | Deep Scan | SD Scan

Antivirus | Security | Privacy | More

Certified Antivirus | Account | TRUSTe Verified

Faked threat report

Click on the buttons

Downloaded phishing app

The Problem

- Enormous effort toward analyzing malicious applications
- App may itself be benign
  - But may lead to malicious content through links
- ***App-web interface***
  - Links inside the app leading to web-content
  - Not well-explored
- Types
  - Advertisements
  - Other links in app

Outline

- App-Web Interface Characteristics
- Solution
- Results
- Conclusion

Outline

- **App-Web Interface Characteristics**
- Solution
- Results
- Conclusion

App-Web Interface Characteristics

- Can be highly dynamic
- A link may recursively redirect to another before leading to a final web page
- Links embedded in apps
  - Can be dynamically generated
  - Can lead to dynamic websites
- Advertisements
  - Ad libraries create links dynamically
  - Ad economics can lead to complex redirection chains

Advertising Overview

ADCOLONY

admob

millennialmedia

友盟 UMENG

INMOBI

Advertisers | Ad networks | Apps / Developers | Users

Ad Networks

- Ad libraries act as the interface between apps and ad network servers
- Ad networks may interface with each other
  - Syndication – One network asks another to fill ad space
  - Ad exchange – Real-time auction of ad space
- App or original ad network may not have control on ads served

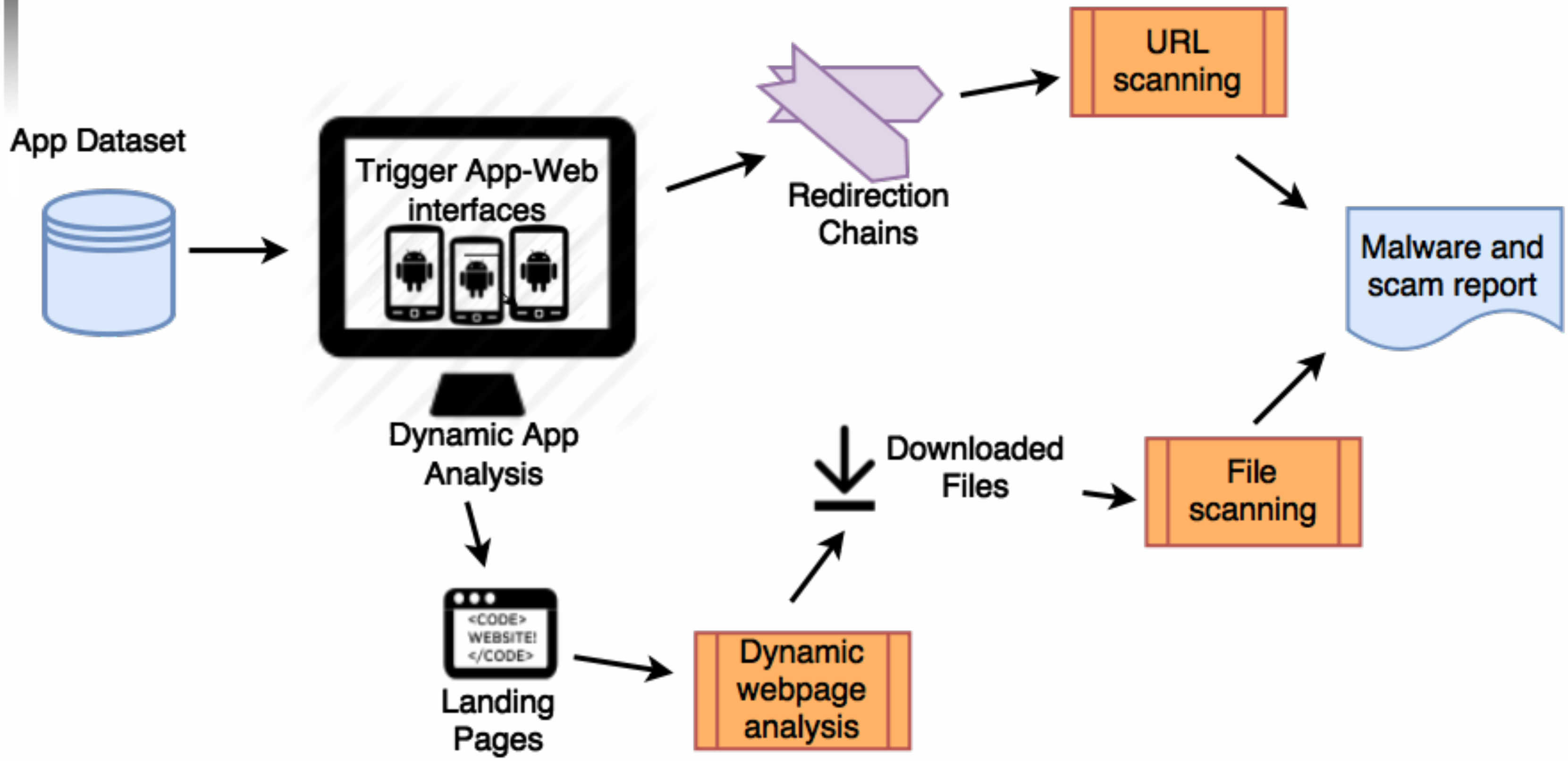
Outline

- App-Web Interface Characteristics
- Solution
- Results
- Conclusion

Solution Components

- **Triggering**: Interact with app to launch web links
- **Detection**: Process the results to identify malicious content
- **Provenance**: Identify the origin of a detected malicious activity
  - Attribute malicious content to domains and ad networks

Solution Architecture



Triggering

- Use AppsPlayground[1]
  - A gray box tool for app UI exploration
  - Extracts features from displayed UI and iteratively generates a UI model

- A novel computer graphics-based algorithm for identifying buttons
  - See widgets and buttons as a human would

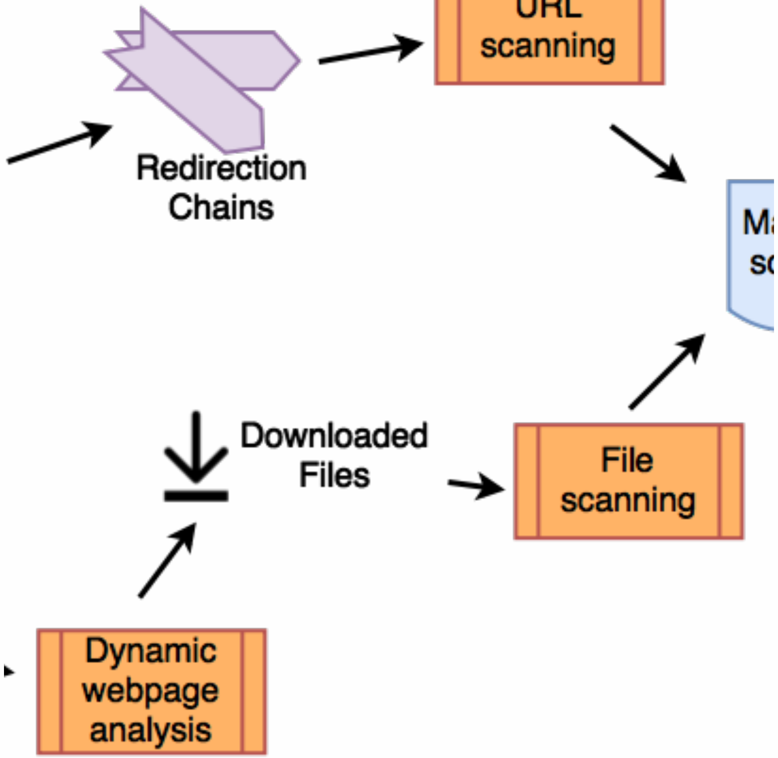
---

[1] Rastogi, Vaibhav, Yan Chen, and William Enck. "AppsPlayground: automatic security analysis of smartphone app
In *Proceedings of the third ACM conference on Data and application security and privacy*, pp. 209-220. ACM, 201

Detection

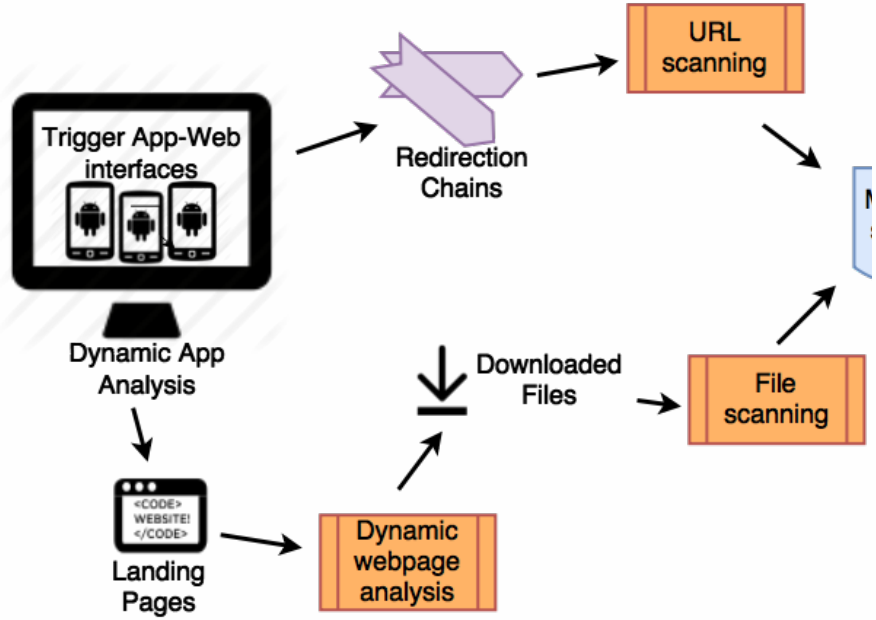
- Automatically download content from landing pages

- Use VirusTotal for detecting malicious files and URLs



Provenance

- How did the user come across an attack?
- Code-level attribution
  - App code
  - Ad libraries
    - **Identified 201 ad libraries**
- Redirection chain-level attribution
  - Which URLs led to attack page or content



Outline

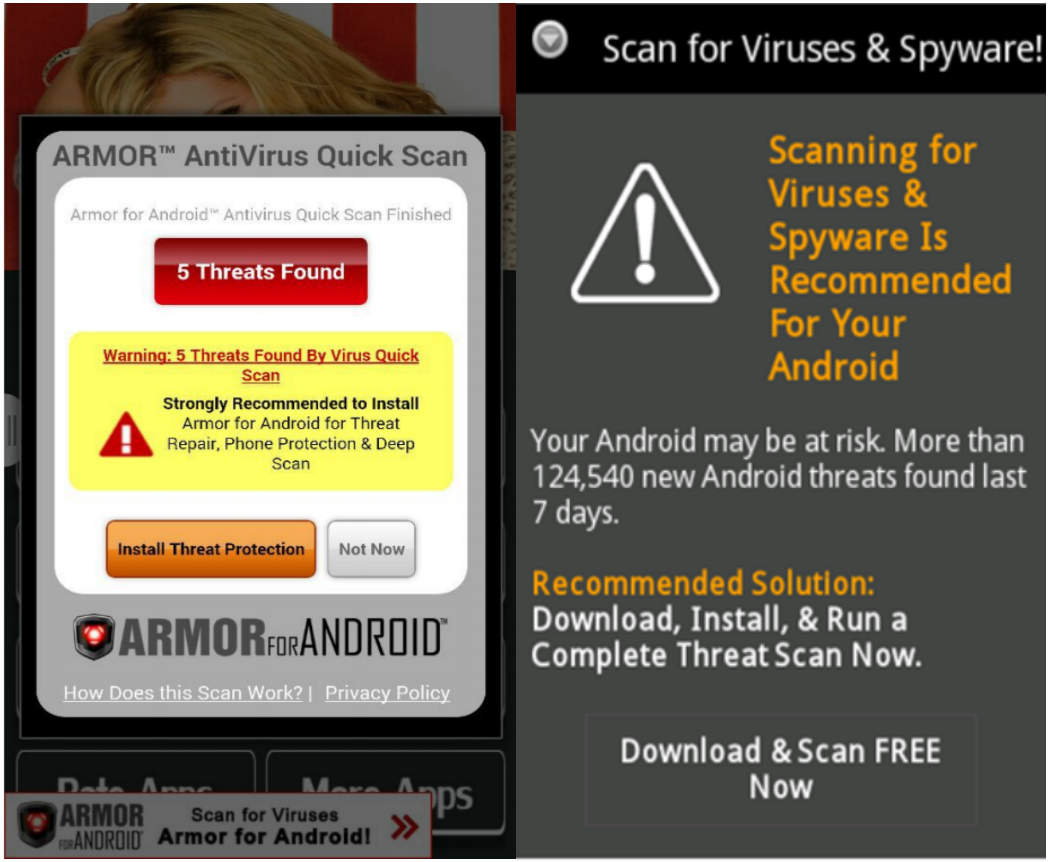
- App-Web Interface Characteristics
- Solution
- **Results**
- Conclusion

Results

- Deployments in US and China
- 600 K apps from Google Play and Chinese stores
  - 91, Anzhi (安智), AppChina(应用汇), Mumayi (木蚂蚁)
- 1.4 M app-web links triggered
- 2,423 malicious URLs
- 706 malicious files

Case Study: Fake AV Scam

- Multiple apps, one ad network: Tapcontext
- Ad network solely serving this scam campaign
- Phishing webpages detected by Google and other URL blacklists about 20 days after we detected first instance



ARMOR™ AntiVirus Quick Scan

Armor for Android™ Antivirus Quick Scan Finished

5 Threats Found

Warning: 5 Threats Found By Virus Quick Scan

Strongly Recommended to Install Armor for Android for Threat Repair, Phone Protection & Deep Scan

Scan for Viruses & Spyware!

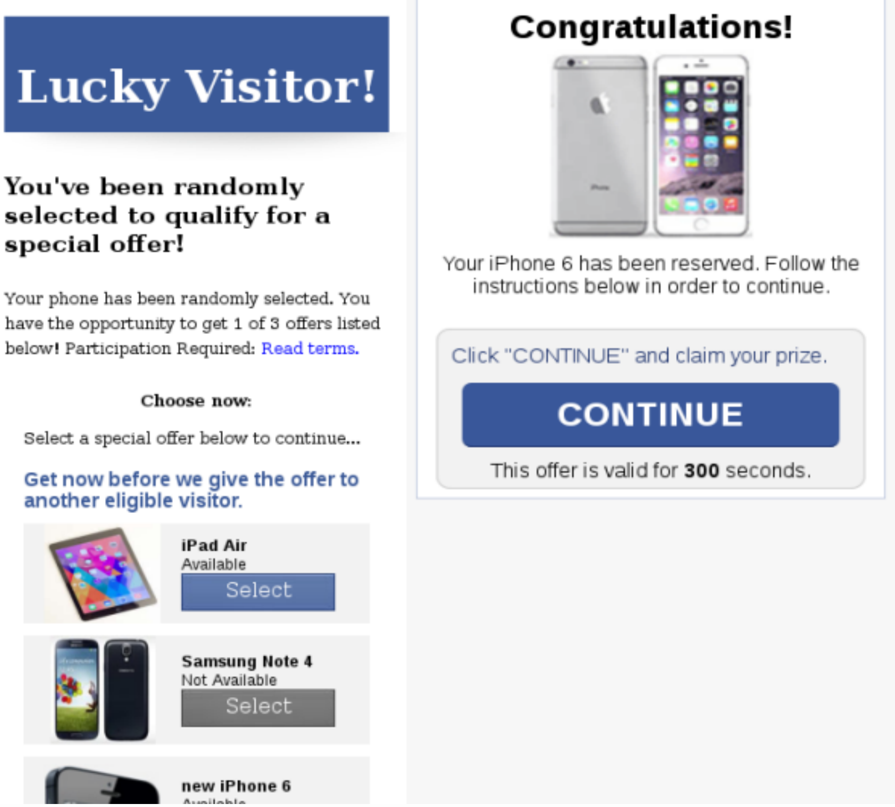
Scanning for Viruses & Spyware Is Recommended For Your Android

Your Android may be at risk. More than 124,540 new Android threats found last 7 days.

Recommended Solution: Download, Install, & Run a Complete Threat Scan Now.

Case Study: Free iPad Scam

- Asked to give personal information without any return
- New email address receiving spam ever since
- Origins at Mobclix and Tapfortap
  - Ad exchanges
  - Neither developers nor the primary ad networks likely aware of this



Lucky Visitor!

You've been randomly selected to qualify for a special offer!

Your phone has been randomly selected. You have the opportunity to get 1 of 3 offers listed below! Participation Required: Read terms.

Choose now:

Select a special offer below to continue...

Get now before we give the offer to another eligible visitor.

Congratulations!

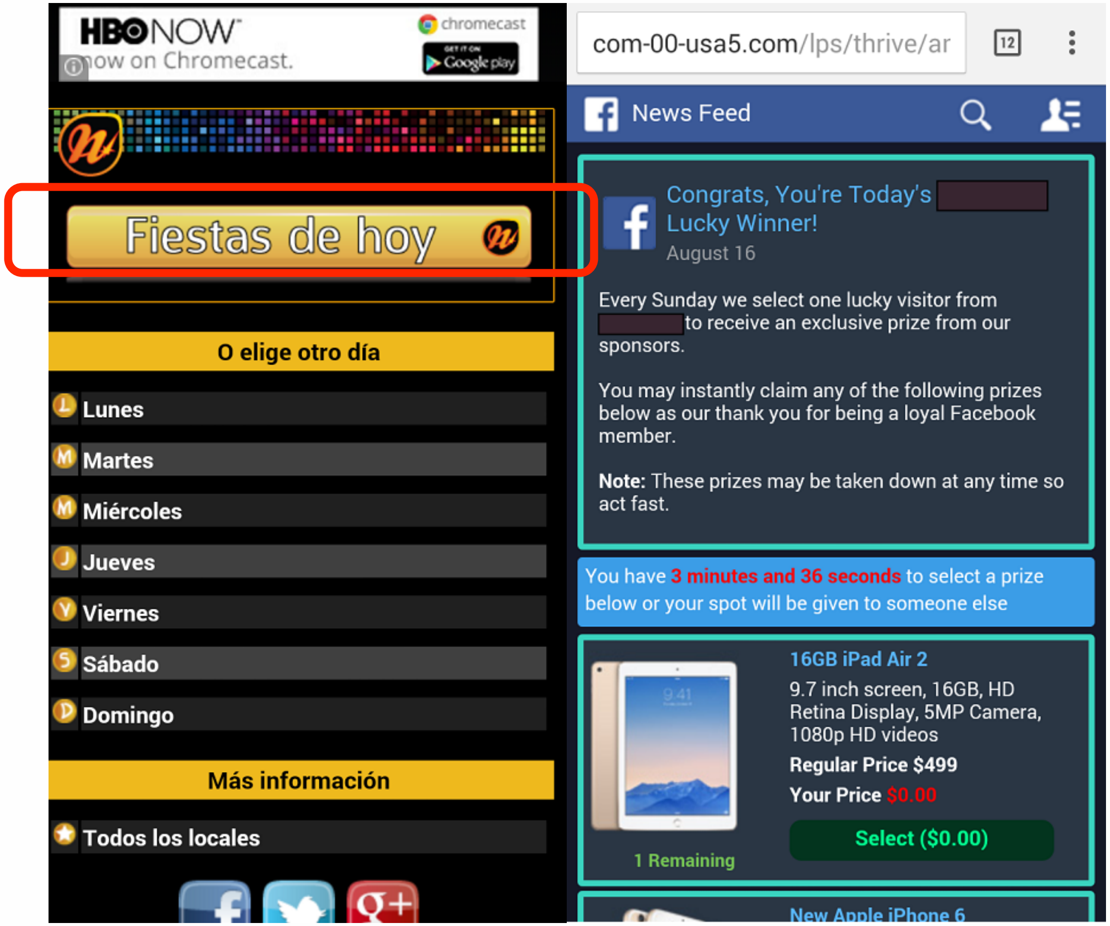
Your iPhone 6 has been reserved. Follow the instructions below in order to continue.

Click "CONTINUE" and claim your prize.

This offer is valid for **300** seconds.

Case Study: iPad Scam from static link

- Another Scam, this time through a static link embedded in app
- Link target opens in browser and redirects to scam
- Not affiliated with Facebook



Congrats, You're Today's [redacted] Lucky Winner!

August 16

Every Sunday we select one lucky visitor from [redacted] to receive an exclusive prize from our sponsors.

You may instantly claim any of the following prizes below as our thank you for being a loyal Facebook member.

Note: These prizes may be taken down at any time so act fast.

You have 3 minutes and 36 seconds to select a prize below or your spot will be given to someone else

16GB iPad Air 2

9.7 inch screen, 16GB, HD Retina Display, 5MP Camera, 1080p HD videos

Regular Price $499

Your Price $0.00

Select ($0.00)

1 Remaining

New Apple iPhone 6

Case Study: SMS Trojan Video Player

給料日まで１日に使える金額

2015/8/3

給料日まであと3日です。

今日は使わない 計算で、

給料日を 含めないで

1日あたり

60,606円使えます。

給料日：6日

残金を入力してください

1 2 1 2 1 2

手机观看【免费APP播

Click on ad

保存

文件名称 mmbo_10483_5080319.apk

文件大小： 3.83M

目录 /storage/sdcard0/Download

- Ad from nobot.co.jp leads to download a movie player
- Player sends SMS messages to a premium number without user consent

Outline

- App-Web Interface Characteristics
- Solution
- Results
- Conclusion

Limitations

- Incomplete detection
  - Antiviruses and URL blacklists are not perfect
  - Our work DroidChameleon[2] shows this

- Incomplete triggering
  - App UI can be very complex
  - May still be sufficient to capture advertisements

[2] Rastogi, Vaibhav, Yan Chen, and Xuxian Jiang. "Catch me if you can: Evaluating android anti-malware against transformation attacks." *Information Forensics and Security, IEEE Transactions on* 9.1 (2014): 99-108.

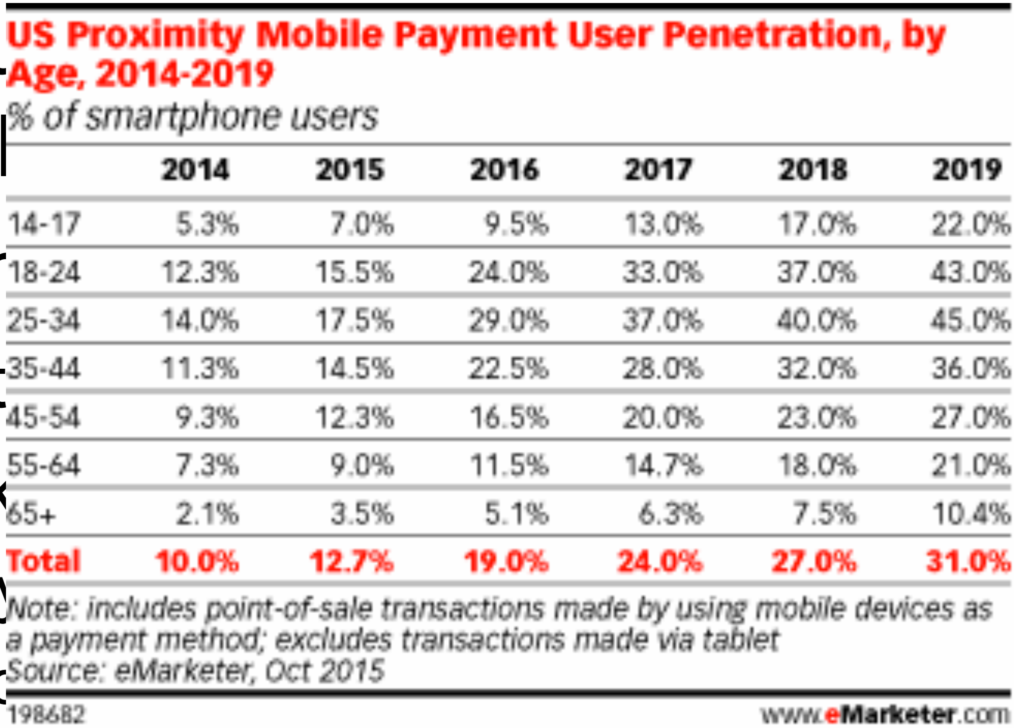
Conclusion and Ongoing Work

- Benign apps can lead to malicious content
- First large scale study to detect malicious ads on Android
- Making it a 24 * 7 service
- Working with ad network providers (e.g., Baidu and Google) and CNCERT for defense
- Only the tip of iceberg, security issues on dynamic code loading (DCL)
  - Detected malware and vulnerabilities that Google Bouncer missed

DROIDCOG: DEVICE-LEVEL MOBILE RISK MANAGEMENT

Motivations

- The growing popularity of mobile payment
- Attack su ser's financial
- Counterr
  - G1: aut
  - G2: risk
    - Heavy 安卓版隐私门)
    - Applic nt detection



US Proximity Mobile Payment User Penetration, by Age, 2014-2019

% of smartphone users

	2014	2015	2016	2017	2018	2019
14-17	5.3%	7.0%	9.5%	13.0%	17.0%	22.0%
18-24	12.3%	15.5%	24.0%	33.0%	37.0%	43.0%
25-34	14.0%	17.5%	29.0%	37.0%	40.0%	45.0%
35-44	11.3%	14.5%	22.5%	28.0%	32.0%	36.0%
45-54	9.3%	12.3%	16.5%	20.0%	23.0%	27.0%
55-64	7.3%	9.0%	11.5%	14.7%	18.0%	21.0%
65+	2.1%	3.5%	5.1%	6.3%	7.5%	10.4%
Total	**10.0%**	**12.7%**	**19.0%**	**24.0%**	**27.0%**	**31.0%**

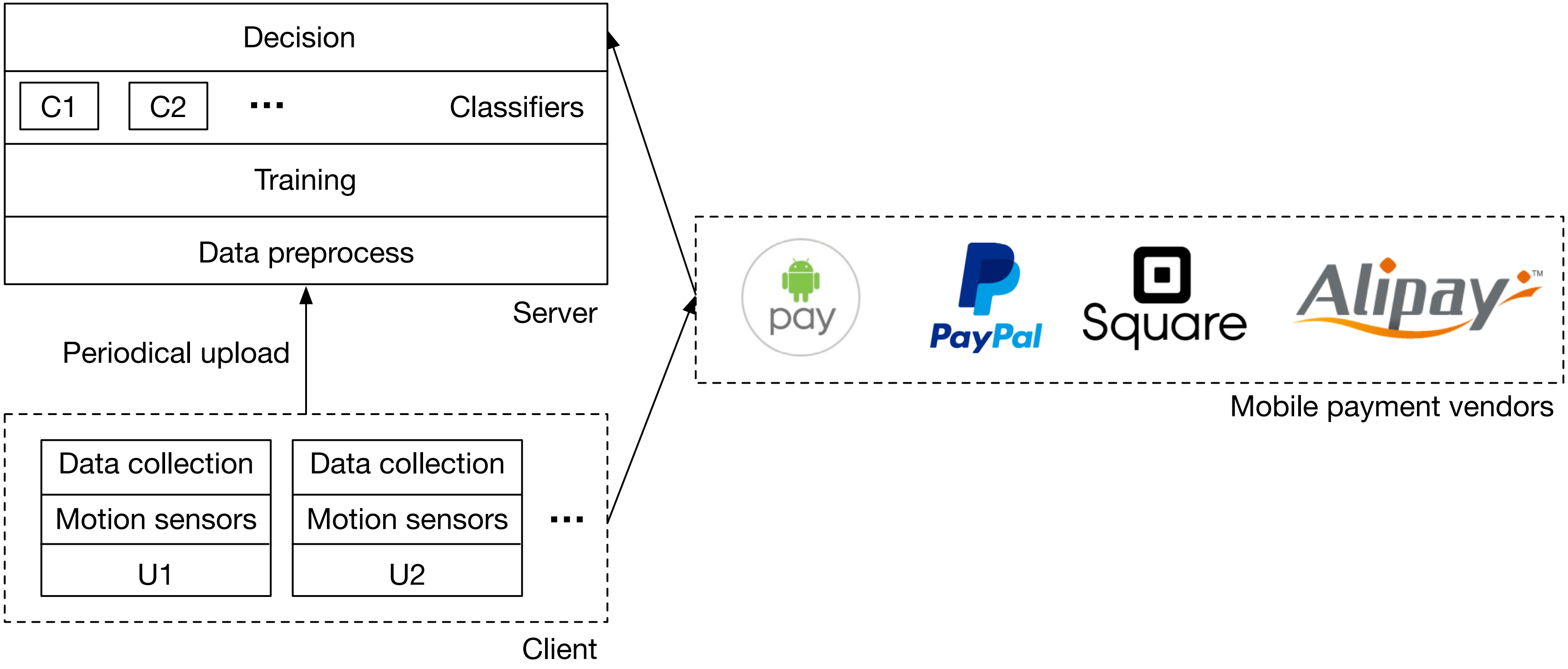
Note: includes point-of-sale transactions made by using mobile devices as a payment method; excludes transactions made via tablet
Source: eMarketer, Oct 2015

198682 www.eMarketer.com

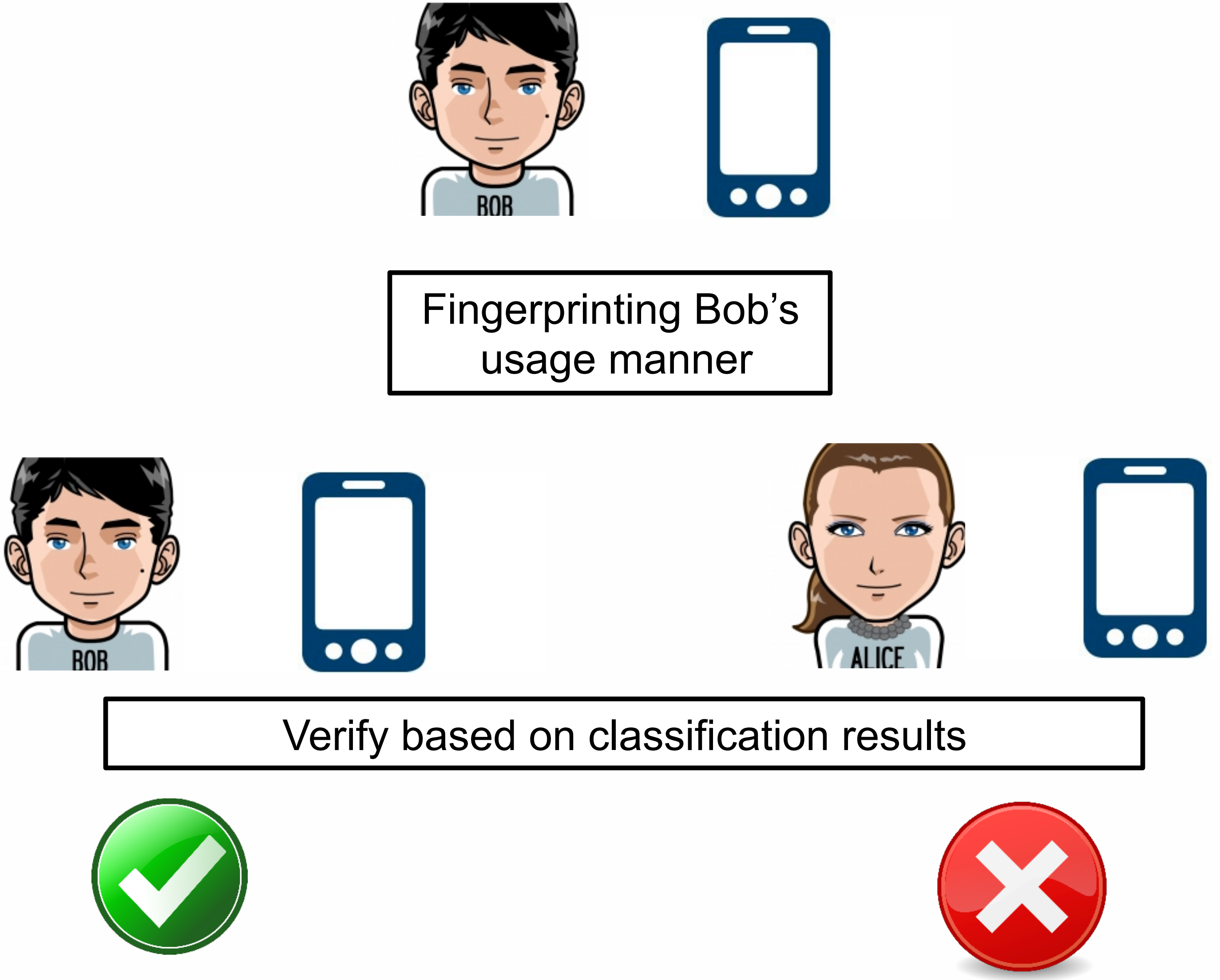
Goal

- A learning-based mechanism for user fraud detection
  - Least user privacy required, high detection accuracy
  - Device-level approach: only one copy of data is uploaded
  - Robust, hard to evade

Goal


Decision
C1
C2
...
Classifiers
Training
Data preprocess
Server
Periodical upload
Data collection
Motion sensors
U1
Data collection
Motion sensors
U2
...
Client
pay
PayPal
Square
Alipay
Mobile payment vendors

Problem Statement



Fingerprinting Bob's usage manner

Verify based on classification results

Challenges

- Lack of features
- Data availability
- Imbalanced dataset
- Noise surrounding
- Unlabeled data

Challenges

- Lack of features
  - Only based on acceleration sensor and gyroscope sensor
  - Feature selection (6 values → 64 features)
- Data availability
- Imbalanced dataset
- Noise surrounding
- Unlabeled data

Challenges

- Lack of features
- Data availability
  - Periodical data collection
  - User activity detection
- Imbalanced dataset
- Noise surrounding
- Unlabeled data

Challenges

- Lack of features
- Data availability
- Imbalanced (classification) dataset
  - Control of distribution of training set
  - Random selection & stratified sampling
- Noise surrounding
- Unlabeled data

Challenges

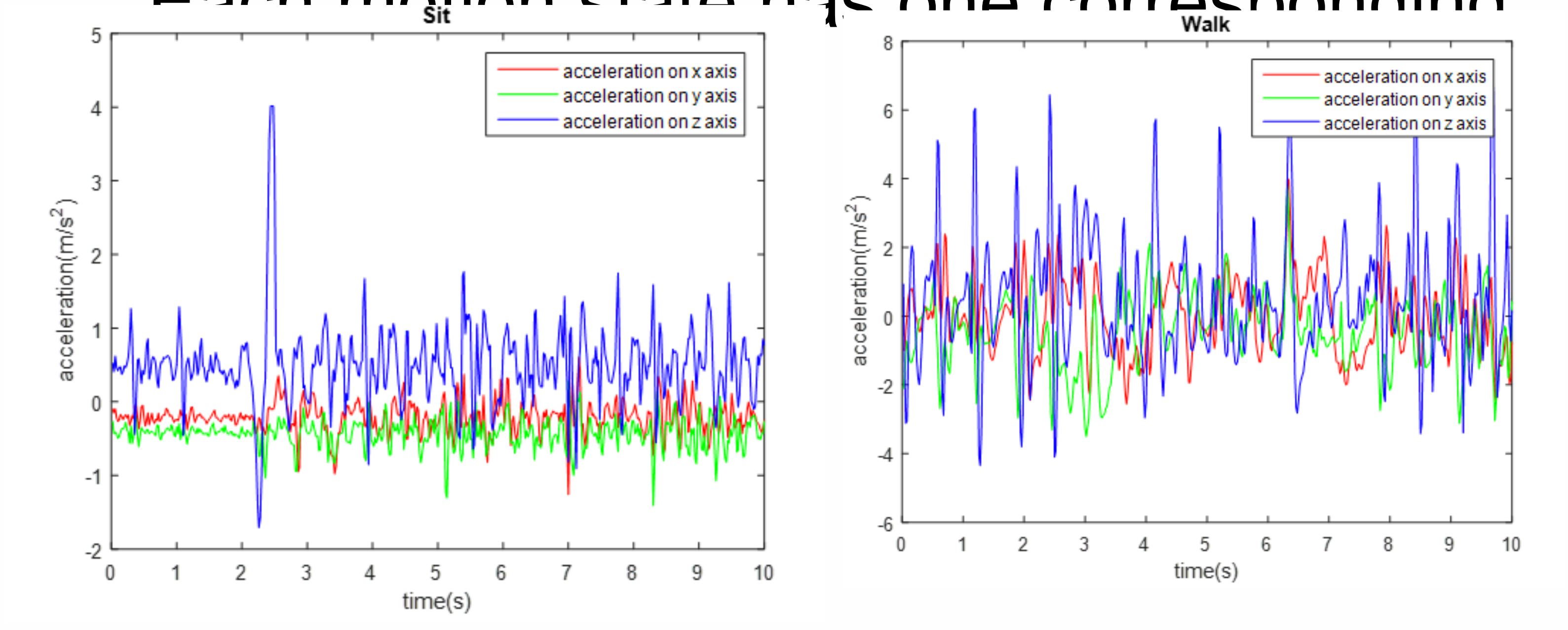
- Lack of features
- Data availability
- Imbalanced dataset
- Noise surrounding
  - Calibrate sensor data based on gravity direction
  - Identify user motion state: static or in motion?
- Unlabeled data

Challenges

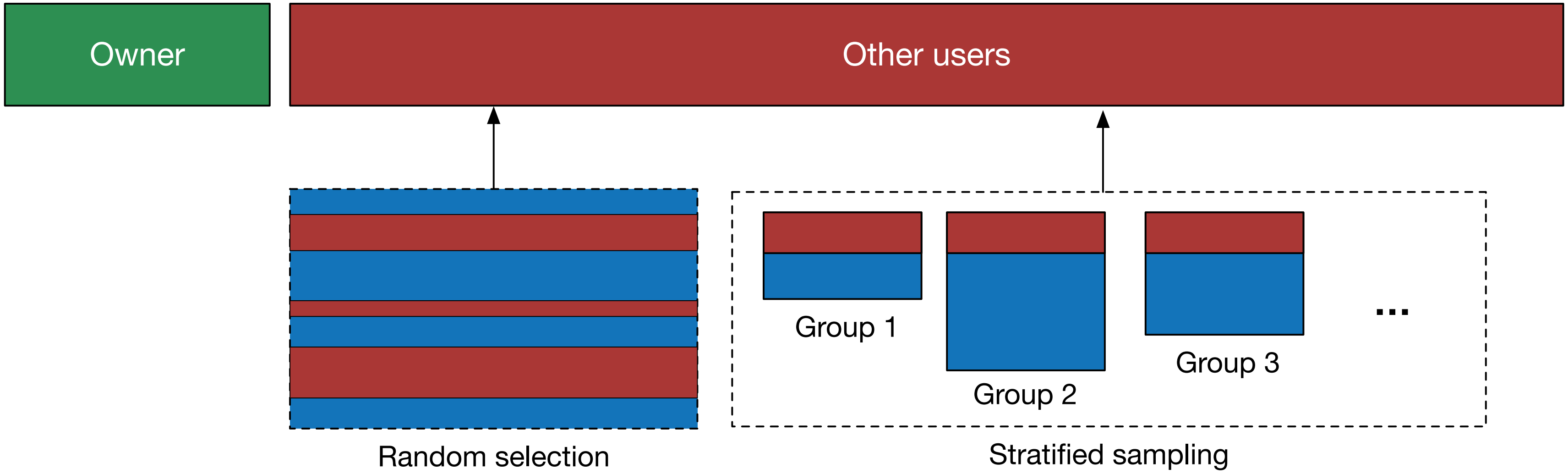
- Lack of features
- Data availability
- Imbalanced dataset
- Noise surrounding
- Unlabeled data
  - Semi-supervised online learning

Data Preprocessing

- Filter useless data on client side
  - The device is put on a flat plane
- Identify motion state on server
  - Each motion state has one corresponding



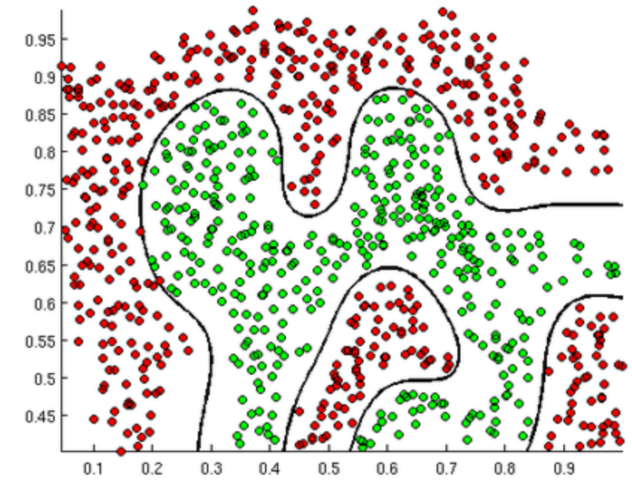
Training Set Construction

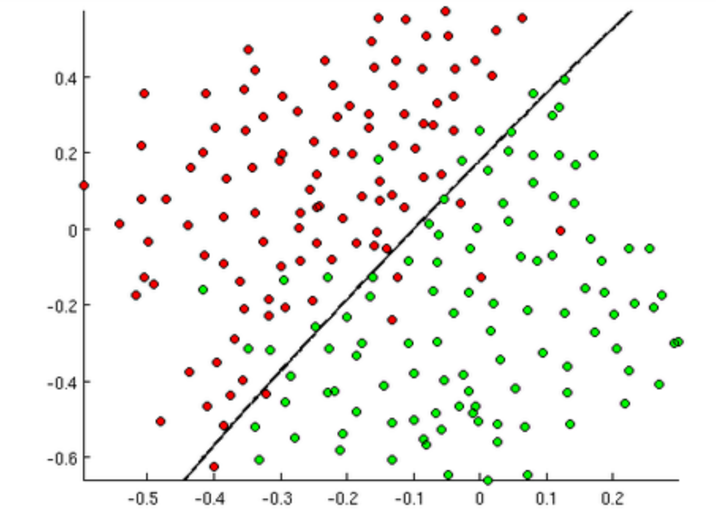


- Random selection v.s. stratified sampling
  - – Similar performance
  - – No cost of grouping user data for random selection

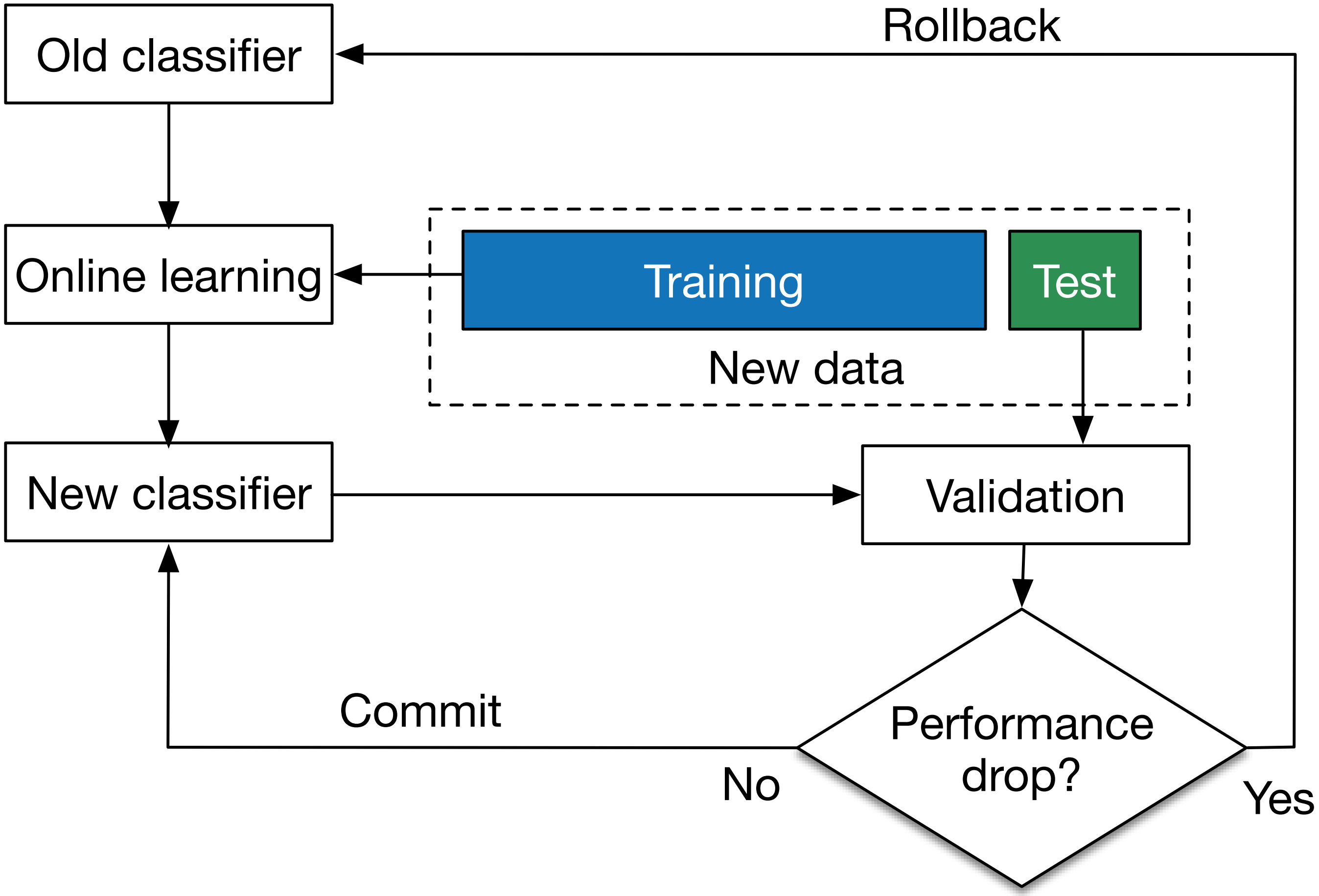
ML Algorithm Selection

- Expectation Maximization (EM): slow
- J48 decision tree: training set over fit, extra cost of tree pruning
- Logistic regression: cannot handle non -linear boundary
- n-li





Semi-supervised Online Learning



Old classifier

Online learning

New classifier

Training

Test

New data

Validation

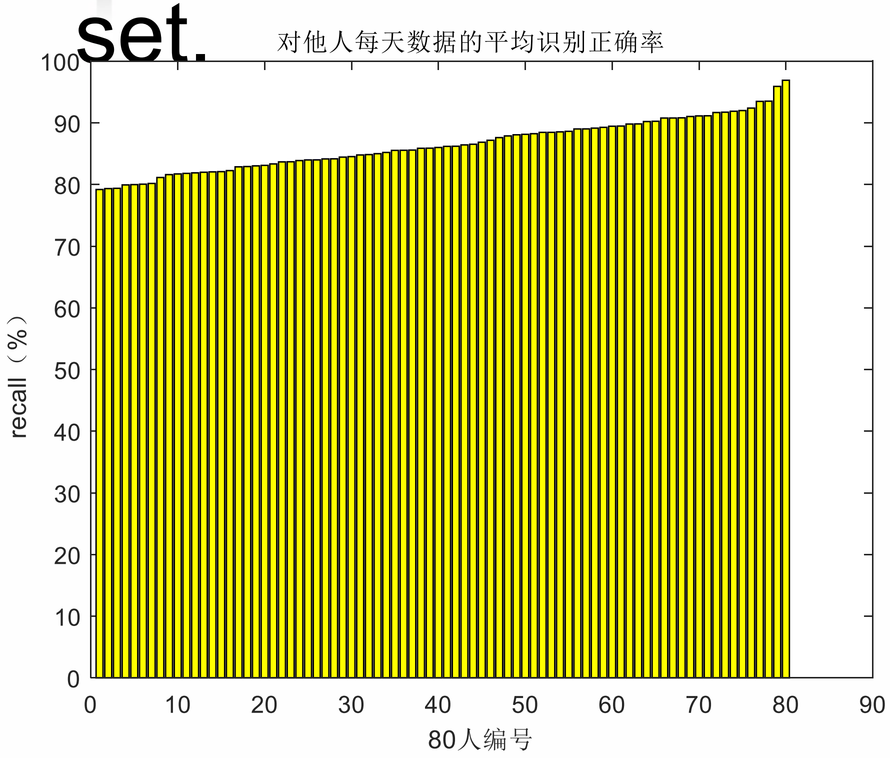
Performance drop?

Preliminary Evaluation

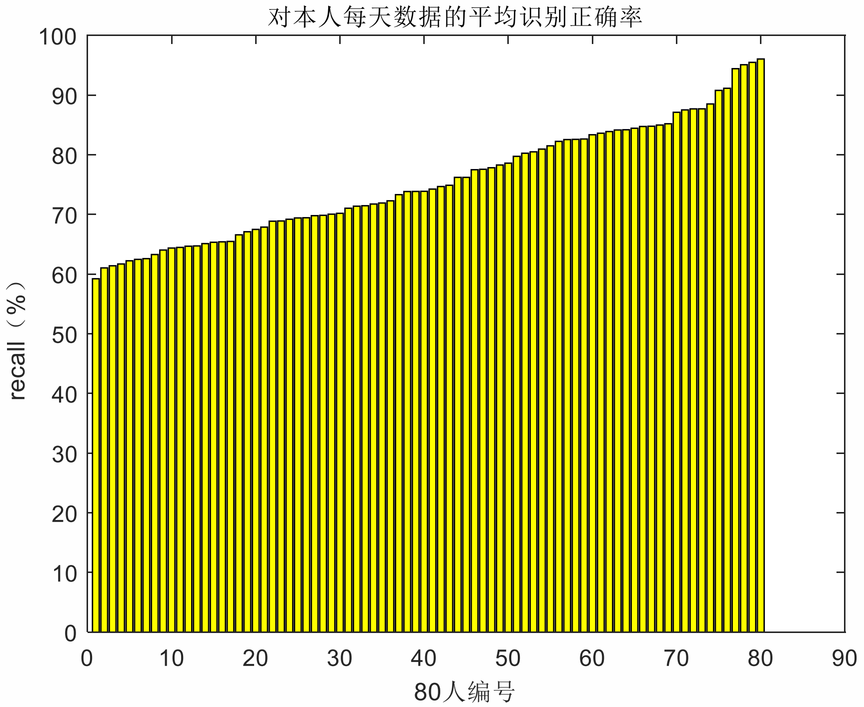
- Data
  - Collected with “Phone manager” (手机管家) by Tencent
  - 1st batch dataset: 210 users
  - 2nd batch dataset: 1516 users
- Metrics
  - Accuracy
    - True positive: owner is correctly identified
    - False positive: other is incorrectly identified as owner
    - False negative: owner is incorrectly identified as other
    - True negative: other is correctly identified
    - R_{owner} = TP/(TP+FN), R_{other} = TN/(TN+FP)
    - ROC curve
  - Overhead
  - Robustness

Accuracy

- 80 users with full data; each user has 4K samples in training set and 1.2K samples in test set.



R_{other}: 86.44%



R_{owner}: 75.50%

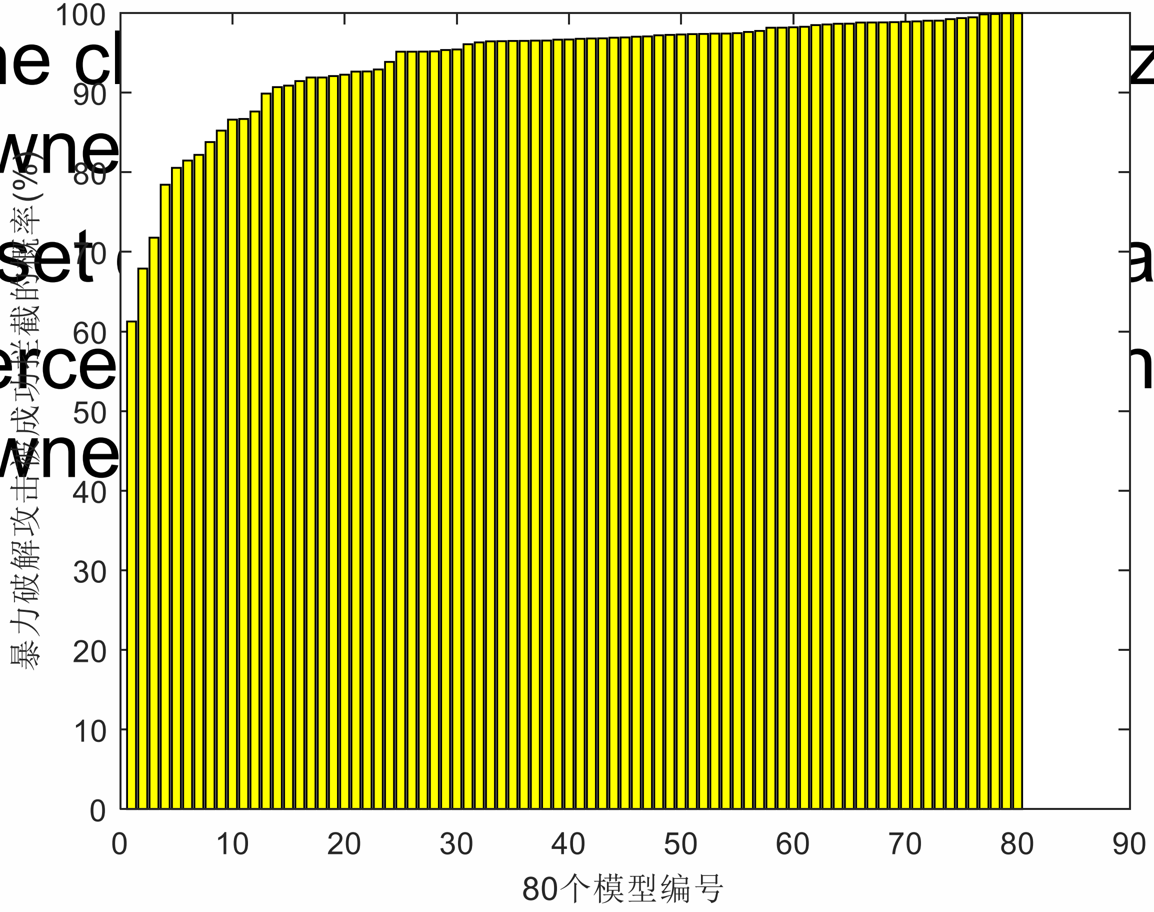
Overhead

- Upload traffic
  - Around 300KB each time, compressed to 90KB.
- Latency (average over 210 users)

#Samples in training set	Training time (s)	#Samples in test set	Test time (s)
13203	18.415	52065	0.639

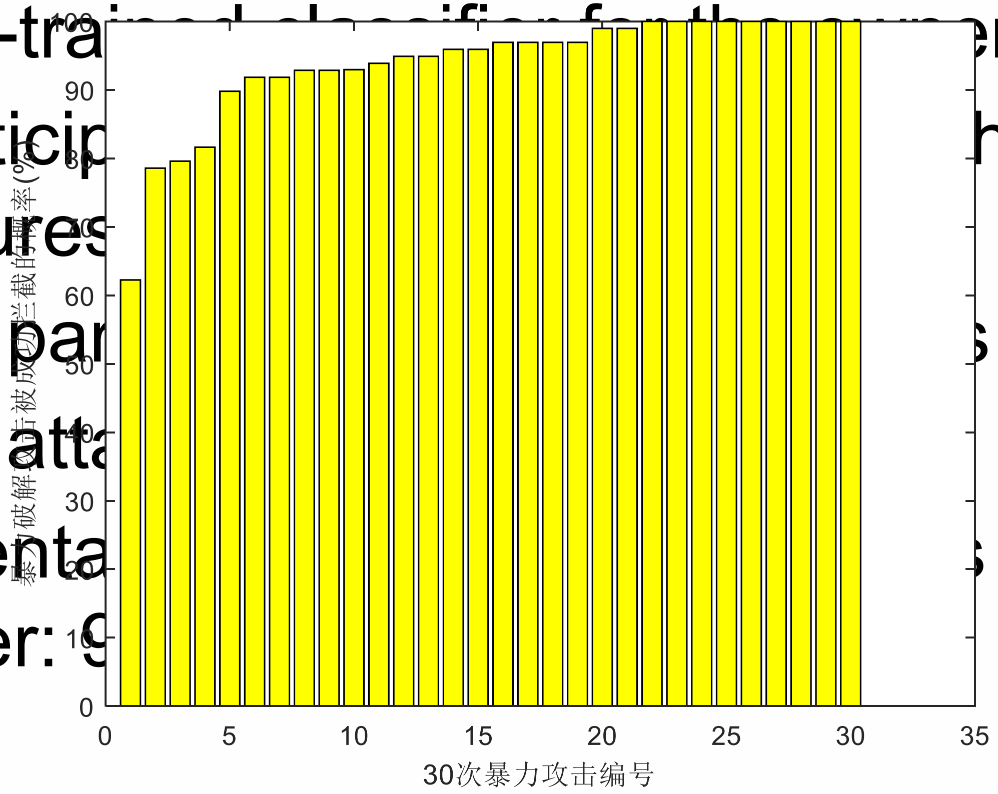
Robustness

- Brute-force attack
  - The cl zed owne
  - A set amples
  - Perce non owne



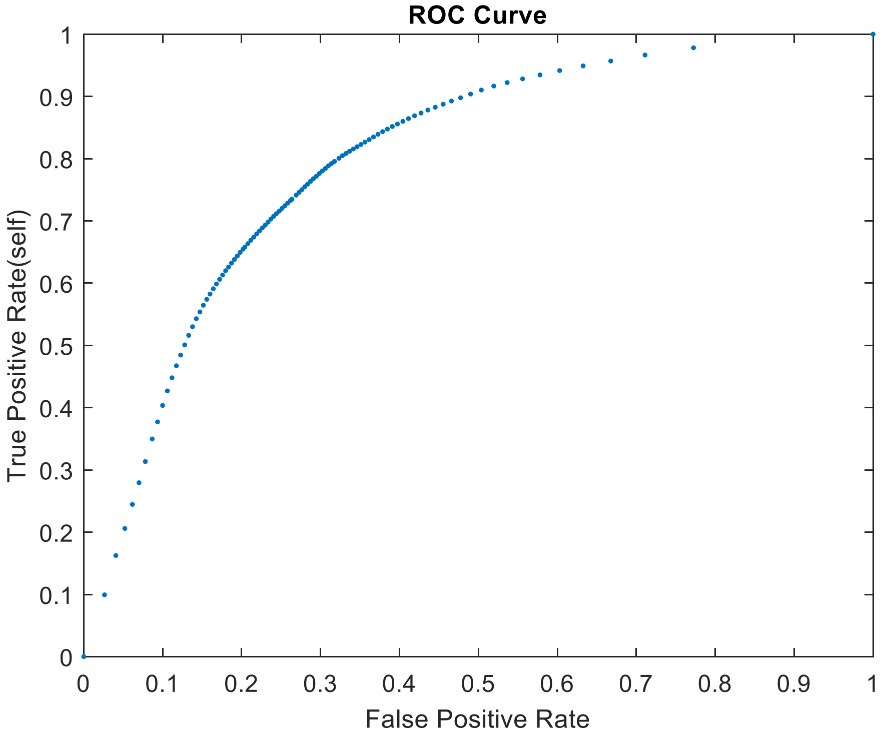
Robustness

- Human attack
  - A pre-tra er
  - 3 particip h various gestures
  - Each par s
  - Each atta
  - Percenta s non owner: 9



ROC Curve

- True positive rate v.s. False positive rate
  - TPR = TP/(TP+FN), FPR = FP/(FP+TN)
  - Changes the classification threshold (0-1)



Conclusion and Ongoing Work

- DroidCog: The first device level user identification system with wild collected sensor data
- Deploy detection system on the phone
- Improve the classification accuracy
  - Explore more usable but privacy insensitive features (e.g. widely used IP address)

- Combine with existing risk management
- Theme of RSA 2016: Connect to Protect

Summary

http://list.cs.northwestern.edu

- Issues for existing mobile anti-virus systems
  - Easy to evade [DroidChamelon]
  - Unable to detect native malware [DroidNative]
  - Unable to detect malware in ads or dynamically loaded content [AdShield]

- Privacy leakage detection and prevention
  - How to find questionable sensitive permissions [AutoCog]
  - Real time tracking & preventing privacy leakage on phone
    - Consumer [PrivacyShield]
    - Enterprise Mobility Management (EMM) [AppShield]

- Fraud detection mostly with app-level risk management [DroidCog]
  - Duplicate detection
  - Privacy infringement