πBox: A Platform for Privacy-Preserving Apps

Sangmin Lee, Edmund L. Wong, Deepak Goel, Mike Dahlin, Vitaly Shmatikov

The University of Texas at Austin
A recent study has found that hundreds of thousands of applications for smartphones steal personal user information and send it to third parties.

At the on-going Black Hat Security Conference in Las Vegas, US-based security firm Lookout revealed the results of its ‘App Genome Project’ report, demonstrating that around 300,000 applications for both Apple’s iPhone and Google’s Android operating systems, were stealing user data.
300,000 Mobile Apps Stealing Personal Data

Path shares photos--oh, and uploads your contacts, too

Summary: The popular photo sharing app is rocked by news that it uploads contacts from iPhone users without permission.
300,000 Mobile Apps Stealing Personal Data

Android Malware Promises Video While Stealing Contacts

Nearly 35% Of Android Apps In China Secretly Steal User Data, Another Sign Of Google’s Lack Of Control

Fake Gmail Android app spies and steals personal information
And because it’s installed without you knowing, you won’t see signs you’ve been bugged
Do you want to install this application?

Allow this application to:

⚠️ **Your messages**
edit SMS or MMS, read SMS or MMS

⚠️ **Your location**
coarse (network-based) location, fine (GPS) location

⚠️ **Network communication**
full Internet access

⚠️ **Storage**
modify/delete SD card contents

⚠️ **Phone calls**
read phone state and identity

Install  Cancel
Do you want to install this application?

Allow this application to:

⚠️ Your messages
edit SMS or MMS, read SMS or MMS

⚠️ Your location
coarse (network-based) location, fine (GPS) location

⚠️ Network communication
full Internet access

⚠️ Storage
modify/delete SD card contents

⚠️ Phone calls
read phone state and identity

Install  Cancel

17% paid attention
From “Android permissions: User attention, comprehension, and behavior.” In SOUPS 2012.
Do you want to install this application?

Allow this application to:

⚠️ **Your messages**
edit SMS or MMS, read SMS or MMS

⚠️ **Your location**
coarse (network-based) location, fine (GPS) location

⚠️ **Network communication**
full Internet access

⚠️ **Storage**
modify/delete SD card contents

⚠️ **Phone calls**
read phone state and identity

[Install] [Cancel]
Do you want to install this application?

Allow this application to:

⚠️ Your messages
edit SMS or MMS, read SMS or MMS

⚠️ Your location
coarse (network-based) location, fine (GPS) location

⚠️ Network communication
full Internet access

⚠️ Storage
modify/delete SD card contents

⚠️ Phone calls
read phone state and identity

Install  Cancel

300,000 app publishers!
Shifting user trust from 300,000 app publishers...
Shifting user trust from 300,000 app publishers... to a few well known brands

Apple  Microsoft  Google  amazon.com
Shifting user trust from 300,000 app publishers...

to a few well known brands that many already trust

Apple
Microsoft
Google
amazon.com
πBox

A platform that allows users to use untrusted apps while providing explicit and useful privacy guarantees
Confine apps for STRONG PRIVACY
aggregate channel

sharing channel
Platform channels for **FUNCTIONALITY**
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of πBox?
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of $\pi$Box?
Per-user, per-app sandbox
Per-user, per-app sandbox
Per-user, per-app sandbox
Per-user, per-app sandbox
Per-user, per-app sandbox spans device and cloud
Per-user, per-app sandbox spans device and cloud
Private vault
read/write

(e.g., settings, search history)

Per-user, per-app sandbox spans device and cloud
Content storage
shared read-only, per-app
(e.g., map data, media)

Private vault
read/write

Per-user, per-app sandbox spans device and cloud
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of πBox?
Just set it and forget it!
The Ronco Showtime Rotisserie Oven
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

releasing true values enable app to signal to publisher
Aggregate channel (shared write only)

uses differential privacy to bound information leak
Aggregate channel (shared write only)

uses differential privacy to bound information leak
Aggregate channel (shared write only)

uses differential privacy to bound information leak
Aggregate channel (shared write only)

uses differential privacy to bound information leak
Aggregate channel (shared write only)

uses differential privacy to bound information leak

see paper for other types of counters (delayed, top-$K$)
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of πBox?
what is shared
when it is shared
with whom it is shared
what is shared
when it is shared
with whom it is shared
what is shared
when it is shared
with whom it is shared
Enter the user ids (separated by commas) that you would like to share with.

This is the data that will be shared:
Dialog box displayed by $\pi$Box
Dialog box displayed by $\pi$Box

$\pi$Box asks whom to share with
Dialog box displayed by $\pi$Box

$\pi$Box asks whom to share with

Users know **when** and **with whom** sharing occurs
Dialog box displayed by $\pi$Box

$\pi$Box asks whom to share with

Users know when and with whom sharing occurs

$\pi$Box confirms content to share
Dialog box displayed by πBox

πBox asks whom to share with

Users know when and with whom sharing occurs

πBox confirms content to share

Users may not know what is shared (steganography)
πBox displays a dialog box asking whom to share with.

Users may not know what is shared (steganography).

πBox confirms content to share.

Difficult for publishers to gain access to private data.

Users know when and with whom sharing occurs.
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of \( \pi \text{Box} \)?
Extended sandbox

Aggregate channel
bounded information leak

Sharing channel
controlled sharing
Extended sandbox
strong confinement

Aggregate channel
bounded information leak

Sharing channel
controlled sharing
Extended sandbox
strong confinement

Aggregate channel
bounded information leak

Sharing channel
controlled sharing
Extended sandbox
strong confinement

Aggregate channel
bounded information leak

Sharing channel
controlled sharing
Extended sandbox
strong confinement
USER WELCOME

NO RISK TO PRIVACY
Extended sandbox
strong confinement

+ 

Aggregate channel
bounded information leak
Extended sandbox
strong confinement

Aggregate channel
bounded information leak

USER GUIDANCE SUGGESTED
MINIMAL RISK TO PRIVACY
Extended sandbox
strong confinement

+ 

Aggregate channel
bounded information leak

+ 

Sharing channel
controlled sharing
Extended sandbox
strong confinement

Counter for ad x

USER STRONGLY CAUTIONED
MAY LEAK INFORMATION WHEN SHARING

Sharing channel
controlled sharing
Outline

How are apps confined within the sandbox?

How does the aggregate channel work?

How does the sharing channel work?

What guarantees are provided to users?

What is the applicability and overhead of $\pi$Box?
Three questions

1. Can real applications benefit from $\pi$Box?
2. How much implementation effort is needed to use $\pi$Box?
3. What is the overhead of using $\pi$Box?
From Google Play (as of Feb. 2013). Based on developer’s description. Core functionality only.

74% of paid apps are green

67% of free apps are yellow
Password Manager

Transcription with feedback

News Reader with ads and sharing
Password Manager

Transcription with feedback

News Reader with ads and sharing
Password Manager

Transcription with feedback

News Reader with ads and sharing
Password Manager

Transcription with feedback

News Reader with ads and sharing
Password Manager

Transcription with feedback

News Reader with ads and sharing

OsmAnd open-source navigation app
changed 174 lines (out of 119,147)

ServStream open-source media streaming app
changed 133 lines (out of 13,193)
Server overheads

**Light workload**

- **With πBox**
- **Without πBox**

**Calculating SHA256 over server-generated 1 MB data**
πBox

Protects users’ privacy from untrusted apps

Provides explicit and simple privacy guarantees
Thank you!