LinkDroid: Reducing Unregulated Aggregation of App-Usage Behaviors

Huan Feng, Kassem Fawaz, Kang G. Shin
Real-Time Computing Laboratory, University of Michigan
An Emerging Threat

Unregulated Aggregation of App-Usage Behaviors

Dynamic Linkability Graph (DLG)

Real-world Evidence

DLG in the real-world

Proposed Solution

LinkDroid: Runtime Monitoring & Mediation
An Emerging Threat

- Political Interests
- Travel History
- Financial Assets
- Dating & Relationship
An Emerging Threat

- Political Interests
- Financial Assets
- Travel History
- Dating & Relationship

Iconography includes logos and icons associated with CNN, Facebook, Snapchat, and other platforms.
An Emerging Threat

- Political Interests
- Travel History
- Financial Assets
- Dating & Relationship

share
leak
An Emerging Threat

Acquisitions of IT Companies

Analytics & Advertising Agency

Surveillance Agency

Icon logos for various companies are included in the diagram.
A curious adversary is able to aggregate usage behaviors of the same user across multiple apps without his knowledge or consent.

The threat of **Unregulated Aggregation of App-Usage Behaviors**

Realistic, financially-motivated, more promising in the future.
A Novel Perspective

An Emerging Threat

Unregulated Aggregation of App-Usage Behaviors

Real-world Evidence

DLG in the real-world

A Novel Perspective

Dynamic Linkability Graph (DLG)

Proposed Solution

LinkDroid: Runtime Monitoring & Mediation
Challenges

Advertising Agency  ?  IT Company

?  ?
A Novel Perspective

Challenges

New paradigms ($\pi$Box, MoRePriv) $\rightarrow$ modify app & ecosystem
A Different Perspective

- Characterize & monitor the **linkability** across mobile apps
  - Two apps are *linkable* if can associate behaviors of the same user
  - Pre-requisites of conducting aggregation

[Diagram showing isolated apps and a Mobile OS]
A Novel Perspective

Mobile OS

OS-Level Information

<table>
<thead>
<tr>
<th>Type</th>
<th>2013-3</th>
<th>2013-10</th>
<th>2014-8</th>
<th>2015-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Android ID</td>
<td>80%</td>
<td>84%</td>
<td>87%</td>
<td>91%</td>
</tr>
<tr>
<td>IMEI</td>
<td>61%</td>
<td>64%</td>
<td>65%</td>
<td>68%</td>
</tr>
<tr>
<td>MAC</td>
<td>28%</td>
<td>42%</td>
<td>51%</td>
<td>55%</td>
</tr>
<tr>
<td>Account</td>
<td>24%</td>
<td>29%</td>
<td>32%</td>
<td>35%</td>
</tr>
<tr>
<td>Contacts</td>
<td>21%</td>
<td>26%</td>
<td>33%</td>
<td>37%</td>
</tr>
</tbody>
</table>

Inter-Process Communications

Explicitly via Binder, or implicitly via shared storage (e.g. SD Card).
A Novel Perspective

isolated

13 Sources of Linkability

<table>
<thead>
<tr>
<th>Category</th>
<th>Type</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS-level Info.</td>
<td>Device</td>
<td>IMEI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Android ID</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAC</td>
</tr>
<tr>
<td></td>
<td>Personal</td>
<td>Phone #</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Account #</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscriber ID</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ICC Serial #</td>
</tr>
<tr>
<td></td>
<td>Contextual</td>
<td>IP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nearby APs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location (PoIs)</td>
</tr>
<tr>
<td>IPC Channel</td>
<td>Explicit</td>
<td>Intent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service Binding</td>
</tr>
<tr>
<td></td>
<td>Implicit</td>
<td>Indirect RW</td>
</tr>
</tbody>
</table>
1. Client-side information is enough
2. Quantify the privacy threat (though upper bound)
Linkable: Two apps are linkable if there exists a path between them.
Linkable: Two apps are linkable if there exists a path between them.
**Linkable**: Two apps are linkable if there exists a path between them.
Linking Ratio ($LR$): # of apps an app is linkable to, divided by all installed apps

Mobile OS

- A (Android ID)
- B (Linking Ratio)
- C
- D
- E
- F
- G
- H (Phone Number)
- I (MAC Address)

Connections:
- A to C
- A to D
- B to D
- C to D
- D to E
- D to F

Intent IPC:
- A to D

MAC Address:
- I

Phone Number:
- G

Linking Ratio ($LR$): # of apps an app is linkable to, divided by all installed apps
**Linking Ratio (LR):** # of apps an app is linkable to, divided by all installed apps

\[ LR_A = \frac{6}{8} \]
Distance: The # of connecting nodes between two linkable apps
Distance: The # of connecting nodes between two linkable apps.

Mobile OS

- **Distance**:
  - The number of connecting nodes between two linkable apps.
**Linking Effort (LE):** average distance between an app and all the apps it’s linkable to.
Linking Effort (LE): average distance between an app and all the apps it’s linkable to

\[ \text{LE}_A = \frac{0 + 2 + 2}{6} \]
Global Linking Ratio (GLR) & Global Linking Effort (GLE)
GLR: Probability of two random apps being linkable (quantity)
GLE: Average distance between two linkable apps (quality)
## Real-world Evidence

### An Emerging Threat

Unregulated Aggregation of App-Usage Behaviors

### A Novel Perspective

Dynamic Linkability Graph (DLG)

### Real-world Evidence

DLG in the real-world

### Proposed Solution

LinkDroid: Runtime Monitoring & Mediation
DLG: A Mobile Extension

• Alternative approaches
  • User-level Interception (Aurasium)
  • Dynamic OS Instrumentation (Xposed Framework)

• Monitor various access to OS-level Info & IPC Channels
  • System Services (Wifi, Telephony, etc)
  • Content Provider
  • Intent Firewall
  • FUSE Daemon
The Alarming Findings

- DLG of 13 users during 47 days using 215 unique apps
  - GLR = 0.81 (two random apps are linkable -> 81%)
  - GLE = 0.2 (control 0.2 additional apps, on average)

- 86% of the apps a user installed are linkable to Facebook, namely his real identity

- Linkability is contributed by various factors (sources)
  - Device ID leads, with others following closely behind
  - Using only contextual information, 40% of apps is linkable to Facebook
Real-world Evidence

Linkability contributed by different sources are proportional to the % of apps accessing each source, except for quasi-identifiers.

- DLG of 13 users during 47 days using 215 unique apps
- GLR = 0.81 (two random apps are linkable > 81%)
- GLE = 0.2 (additional apps, on average)
- 86% of the apps a user installed are linkable to Facebook, namely his real identity
- Linkability is contributed by various factors (sources)
  - Device ID leads, with others following closely behind
  - Using only contextual information, 40% of apps is linkable to Facebook
Functional Analysis

• **OS-level Information**
  - *Device ID* no need for the actual identifiers
  - *Personal ID* abuse user accounts & phone #
  - *Contextual ID* exploit Location & nearby AP

• **IPC Communications**
  - Apps report their installation using *Intents* (WeChat)
  - Apps *bind to service* & exchange user IDs (Facebook, AdMob)
  - Apps read identifiers written by other apps (Qingting Radio)

• Subject to personal preference and application context
Proposed Solution

- An Emerging Threat
  Unregulated Aggregation of App-Usage Behaviors

- A Novel Perspective
  Dynamic Linkability Graph (DLG)

- Real-world Evidence
  DLG in the real-world

- Proposed Solution
  LinkDroid: Runtime Monitoring & Mediation
LinkDroid

• Designed with **practicality** in mind
  • No modification of apps, no additional trusted parties
  • Works purely on the client-side

• A **new** dimension to privacy protection on mobile OS
  • How app behaviors implicitly affect linkability
  • Opt-out & reduce unnecessary links

• Features provided by LinkDroid
  • Install-time Obfuscation
  • DLG-powered Runtime Monitoring
  • Unlinkable Mode
Install-time Obfuscation

Device ID (Android ID, IMEI, etc)
Runtime Monitoring
Descriptive Risk Indicator

Quantitative Risk Indicator

Opt-out Options

Behavioral Description of the Scenario
Unlinkable Mode

A new instance installed on a new device
Design of LinkDroid
Evaluation

• Evaluated on the same set of 13 participants as in the measurement
  • Replay traces collected in the measurement with LinkDroid features

• GLR (two random apps being linkable): 81% to 21%

• GLE (additional apps required to link two apps): 0.22 to 0.68
  • Under most scenarios, at least one additional app is required

• Apps directly linkable to Facebook dropped from 86% to 18%
DLG of a representative user *before* and *after* applying LinkDroid. (Red circle is the Facebook app)
Takeaway

Leaked (shared) information should **NOT** be linkable unless **REALLY** necessary

Linkability: a useful but **MISSING** notion in the mobile ecosystem

Anonymous (unlinkable) in-app behaviors should be a **BASIC** right
Questions?

LinkDroid: Reducing Unregulated Aggregation of App-Usage Behaviors

Huan Feng
huanfeng@umich.edu