

Detecting Privacy Leaks in the RATP App: how we proceeded and what we found

Jagdish Prasad Achara ^{Speaker},
James-Douglass Lefruit, Vincent Roca, Claude Castelluccia

INRIA Rhone-Alpes

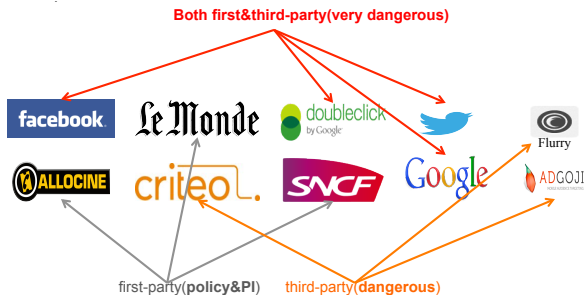
November 15, 2013

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App
- 4 Conclusions
- 5 References

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App
- 4 Conclusions
- 5 References

A large number of actors present on Smartphones

- ▶ Due to revolutionary arrival of AppStore model of App distribution
- ▶ These actors could be categorized as follows:
 - ▶ First-party : whose services are used by the user explicitly (App owners, OS provider, Cellular (GSM/CDMA) service providers etc.)
 - ▶ Second-party : the user himself
 - ▶ Third-party : to whom the user doesn't directly interact with (Advertisers, Analytics companies, performance monitors, crash reporters, push senders etc.)



3

Difficult to trust all these actors

- 1 Various scandals in the past
 - ▶ e.g. Twitter and Path uploading users' contacts to their servers [4], [8]
- 2 WSJ: What They Know - Mobile [9]



Categories of data:

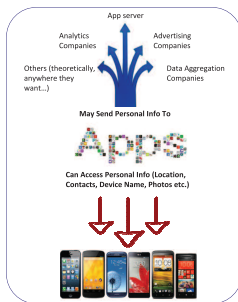
User, password
Contacts
Age, gender
Location
Phone ID
Phone number

App name	iPhone		Android			
	Username, Password	Contacts	Age, Gender	Location	Phone ID	Phone number
0.03 Seconds Pro	Does not transmit data	Does not transmit data	Does not transmit data	Does not transmit data	Transmits data to third parties	Does not transmit data
Age My Face	Does not transmit data	Does not transmit data	Does not transmit data	Does not transmit data	Transmits data to third parties	Does not transmit data
Angry Birds	Transmits data to app owner	Transmits data to app owner	Does not transmit data	Transmits data to app owner	Transmits data to third parties	Does not transmit data
Angry Birds Lite	Transmits data to app owner	Does not transmit data	Does not transmit data	Transmits data to app owner	Transmits data to third parties	Does not transmit data
Aurora Feint II: Lite	Does not transmit data	Does not transmit data	Does not transmit data	Transmits data to app owner	Transmits data to third parties	Does not transmit data
Barcode Scanner (BahnTech)	Does not transmit data	Does not transmit data	Does not transmit data	Does not transmit data	Transmits data to third parties	Does not transmit data

Legend:
Does not transmit data (white)
Transmits data to app owner (blue with vertical lines)
Transmits data to third parties (solid blue)

More opportunities for personal information (PI) leakage on Smartphones

- 1 **Not limited to web browsers** as is the case mostly in desktops/laptops
- 2 App code (coming from different parties) runs on the device
- 3 They **contain a lot of info about user interests and behaviors**
 - ▶ Sensors (GPS, Camera) and Comm. chips (WiFi, GSM/CDMA) generate PI
 - ▶ At the center of our cyber activities, and very personal (not shared usually)
 - ▶ Almost all-time Internet connectivity and barely turned-off



This leads to **detailed and accurate user profiling**

Smartphones are well-suited to Marketers

- ▶ A direct consequence is **the large presence of online advertisers/trackers**

admob ⁽⁽⁽⁽⁽)



doubleclick
by Google



Flurry

criteo

and many others...

There is a clear need for **"tracking the trackers"**

RATP: French public company that is managing the Paris subway (metro)

- ▶ It provides very useful App for both Android and iOS helping users to easily navigate in the city.
- ▶ We found RATP App leaking user personal information (PI) in total contradiction to their In-App privacy policy.

- ▶ This talk details/discusses
 - ▶ **the Methodology** (a combination of static and dynamic analysis techniques) we used **to analyze both Android and iOS Apps of RATP.**
 - ▶ **our Findings w.r.t. PI leakage by RATP Apps:** what user PI is leaked and to which servers...
 - ▶ **the responsibilities of various actors** (OS providers and App developers), in general, to stop the practice of user PI leakage.

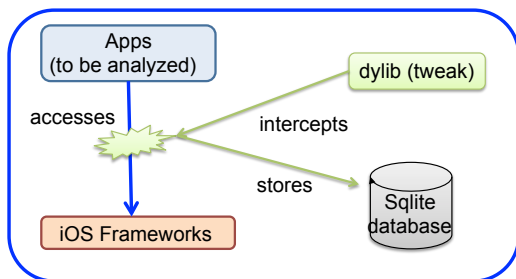
Outline

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App
- 4 Conclusions
- 5 References

Instrumented version of iOS

To detect private data leakage

- ▶ Code in a dynamic library that is loaded in Apps to be analyzed at runtime¹
 - ▶ Using Objective-C runtime
 - ▶ Replacing C/C++ functions at assembly level



¹MobileSubstrate [2] and Theos [7] simplify this task!

Privacy Leaks to Adgoji

Adgoji: A mobile audience targetting company

Listing 1.1. Data sent through SSL by iOS App of RATP (Instance 1)

```
UTF8StringOfDataSentThroughSSL = { "p": { "kernel_task", "launchd",  
  "UserEventAgent", "shsettingsd", "wifi", "powerd", "lockdown",  
  "mediaserverd", "adIDASpender", "locationd", "imgent", "iagtrnsupportd",  
  "fsventd", "fairplayd", "configd", "rsbd", "ComOnCenter", "RTServer",  
  "notifiyd", "aggregated", "networkd", "Timesasstored", "apsad", "MyWiCore",  
  "diatnoted", "cccd", "filecoordination", "installd", "absintbed", "timed",  
  "geod", "networkd_privile", "isd", "spcd", "accountd", "notification_pro",  
  "corwaymbolicatio", "assetd", "AppleIDAuthAgent", "dataaccessd",  
  "SOSlipes", "backboardd", "sysd", "syslogd", "dshorag", "SpringBoard",  
  "Facebook", "iFile", "iMessage", "MobilePhone", "MobileSIP2P",  
  "MobileSMServi", "webbookmarkd", "napollient", "mobile_installd",  
  "AppStore", "syncdefaults", "social", "remd", "RATP", "pasteboardd",  
  "additional": { "device_language": "en", "country_code": "FR",  
    "adgoji_sdk_version": "v2.0.2", "device_system_name": "iPhone  
OS", "device_jailbroken": true, "bundle_version": "5.4.1",  
    "vendorid": "CE0C8023-96A2-4005-A1FB-96E3CDA1E79", "allows_voip": false,  
    "device_model": "iPhone", "wscsd4-xyz": "00facd10c20", "wscsd4":  
    "49e8a601-9753-108c-45c9-b417383742d", "bundle_identifier":  
    "com.ratp.ratp", "system_os_version_name": "iPhone OS", "device_name":  
    "2jy3k8ib v iPhone", "bundle_executable": "RATP",  
    "device_localized_model": "iPhone", "sysudid":  
    "7679d76a7037f4d48548a8c8c979ad8c8c8d", "s":  
    {"78E4FAA-FF62-48EF-B619-21145CF16564": {"m": "start",  
      "e": "1369926018", "name": "ERx8Ac"} } } }
```

Listing 1.2. Data sent through SSL by iOS App of RATP (Instance 2)

```
UTF8StringOfDataSentThroughSSL = { "s": [ "fb210831918949520",  
  "fb108880882526064", "evernote", "fbauth2", "fbauth", "fb", "fblogin",  
  "fspot-image", "fb308918024569", "fspot", "fsq",  
  "pjq45qactoi.jhuqf5121d5tyur0zosvmfadyw0pvd4b434e+authorize",  
  "fsq+pjq45qactoi.jhuqf5121d5tyur0zosvmfadyw0pvd4b434e+reply",  
  "fsq+pjq45qactoi.jhuqf5121d5tyur0zosvmfadyw0pvd4b434e+post",  
  "foursquareplugins", "foursquare", "fb86734274142", "fb124024574287414",  
  "instagram", "fsq+kylm3gjcvtswk4rambrt4uyzq1dqoc0n2hyjgcvbcb54rj+post",  
  "fb-messenger", "fb237759909591655", "RunKeeperPro", "fb62572192129",  
  "fb76446685859", "fb142349171124", "soundcloud", "fb19507961798",  
  "x-soundcloud", "fb110144382383802", "mailto", "spotify", "fb134519659678",  
  "fb174829003346", "fb109306535771", "tj459035295", "twitter",  
  "com.twitter.twitter-iphone", "com.twitter.twitter-iphone+1.0.0",  
  "fweetie", "com.atubits.Tweetie2", "com.atubits.Tweetie2+2.0.0",  
  "com.atubits.Tweetie2+2.1.0", "com.atubits.Tweetie2+2.1.1",  
  "com.atubits.Tweetie2+3.0.0", "FTP", "PPClient", "fb184136951108"] }
```

The user PI sent is

- ▶ WiFi MAC Address
- ▶ List of currently running processes
- ▶ Device Name
- ▶ OpenUDID
- ▶ Advertising ID
- ▶ List of URLSchemes available on the device (to know if corresponding Apps are installed)

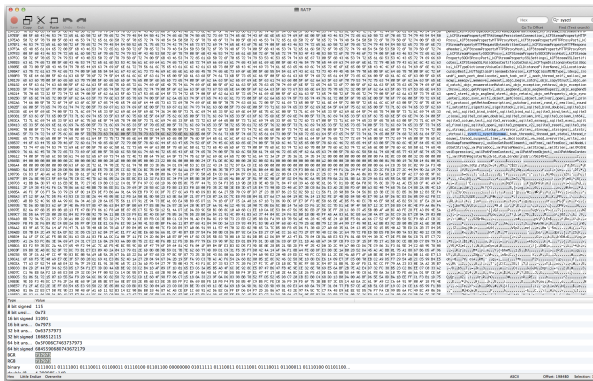
Adgoji: how to know Apps installed on the device

It is very useful info to infer the user interests and behaviour.

- ▶ No API provided by iOS Frameworks to do so...

But techniques exist to know a subset of Apps installed (if not all!)

- ▶ Use of 1) sysctl function (in libc) and 2) URLScheme class [1].

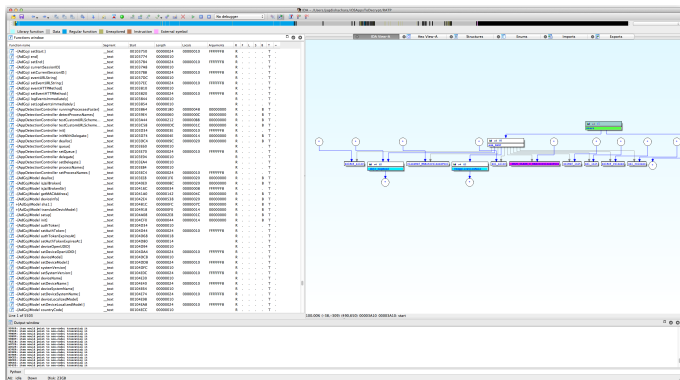


Presence of "sysctl" String in decrypted App binary confirms its use in the code written by the App developer



Adgoji: confirmation of its presence in the App (1)

RATP iOS App binary opened in IDA after decryption



Privacy Leaks to Sofialys company

Sofialys: A mobile advertising company [5]

Listing 1.3. Data sent by iOS App of RATP in cleartext

```
UTFStringOfDataSentInCLEAR = {"uage":"","confirm":"1", "imei":  
"9c7a916a17037454ed05debc8c3e977edbc0bcdd", "osversion":"iPhone6.1.2",  
"odin":"1b84e4efaf650cb9a264a2ff23ca7a67b9bd72f6", "umail":"","  
"carrier":""," "user_position": "45.218156;5.807636", "long":"","  
"ua":"Mozilla/5.0 (iPhone; CPUiPhoneOS6_1_2likeMacOSX) AppleWebKit  
/536.26 (KHTML, likeGecko) Mobile/10B146", "fingerprint":{"v1":  
{"i":"3739335834508445""b""c5kkekiLx11ghUfu3Ht43bUZwCHHBNbRO  
9A04it+wtPPCBJagCio7tgBdM1q6T244EwHnKRzeh1ybrMhKy2SztEU5tD5u5Q  
7HAisR57BYIun9aQdp0NsXwp7BXhohS92daScYcMDALqKqYKZDriEjQW  
wtjvR9MrIKfE52EwNcA9CJkUIT9q7sXkqkvalo0M7tMrNdMiIQYyH0tdNJ+  
ax7Ujau/IQ4pPasSXk/m6BIFsAFhjF0ng0NuSwtL7e7r95s8wQhWy+  
EvJUCHPivIRXZYldCbjfdkrkvNgHZcH59Fj0dBz9Ugbyoj4a/Z60S1U+  
EatvNswORMQqdE8djvJmXkGCmwoheU10uQatr4ppA="}}, "ugender":"","  
"os":"iPhone", "adid": "496EA6D1-5753-40B2-A5C9-5841738374A2",  
"uphone":""," "sdkversion":"5.0.3", "test":""," "lat":""," "udob":"","  
"pid":"4ed37f3f20b4f", "lang":"fr_FR", "network":"wifi",  
"time":"2013-05-3015:45:04", "alid":"186", "sal":""," "uzip":""}
```

The user PI sent is (IN CLEAR-TEXT):

- ▶ The exact user location
- ▶ Advertising ID and UDID

Responsibility of Apple

AdID: An alphanumeric string unique to each device, used only for serving advertisements.

But Advertising Identifier only gives an illusion to the user that he is able to opt-out from device tracking:

- ▶ WiFiMAC Address (using sysctl function in libc library)
- ▶ Using UIPasteboard to generate a unique identifier across the device
- ▶ Device Name
- ▶ UDID (still being used by "old" Apps even if deprecated)

Apple privacy dashboard is not enough because:

- ▶ A&A libraries included by the App developer have access to the same set of user's private data as the App itself.
- ▶ Behavioral analysis is required.
- ▶ Finer granularity permission is needed

Apple cannot ignore this trend.

Outline

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App**
- 4 Conclusions
- 5 References

Instrumented version of Android

- ▶ We use Taintdroid [6] to track user PI flow (and hence the leakage of PI over network)
- ▶ We also change the source code of Android itself (only the APIs of our interest e.g. network APIs to look for the data sent over the network) to fill-in the gaps
 - ▶ Taintdroid could miss the leakage of some PI [3]
 - ▶ Some PI (e.g. Android ID) can't be tainted due to false positives.
- ▶ In addition, we also use static analysis to confirm some observations.

Privacy Leaks to Sofialys company

Listing 1.4. Data sent in cleartext by Android App of RATP

```
DataSentInCLEAR =  
{ "user_position": "45.2115529;5.8037135", "ugender": "",  
  "test": "", "uage": "0", "imei": "56b4153b8bd2f6fd242d84b3f63e287", "napp":  
  null, "uemail": "", "pid": "4ed37f3f20b4f", "alid": "114", "uzip": "",  
  "osversion": "3.0.31-g396c4dfdirty", "lang": "en_En", "sal": "", "network":  
  "na", "adpos": null, "time": "Tue Jun 04 12:05:39 UTC+02:00 2013",  
  "sdkversion": "3.2", "ua": "Mozilla/5.0(Linux; U; Android 4.1.1;  
  fr-fr; Full AOSP on Maguro Build/JR003R) AppleWebKit/534.30  
  (KHTML, like Gecko) Version/4.0 Mobile Safari/534.30", "udob": "",  
  "carrier": "Orange F", "longitude": "0.0", "latitude": "0.0",  
  "freespace": null, "unick": null}]
```

The user PI sent is:

- ▶ The exact location of the user
- ▶ the MD5 hash of the device IMEI
- ▶ the SIM card's carrier/operator name

Is the hashing of IMEI sufficient to guarantee anonymity?

```
~/doc/oc/Hashcat-plus-0.14/benchFinalGPU more PREFIX_IMEI
W*4?5?1?5?
~/doc/oc/Hashcat-plus-0.14/benchFinalGPU ./mp12_bin 7d7d7d7d7d | .../cudaHashcat-
plus12_bin -m 0 hash MD5_IMEI -r PREFIX_IMEI -r theRule2digit
cudaHashcat-plus v0.14 by atom starting...

Hashes: 1 total, 1 unique salts, 1 unique digests
Bitmaps: 8 bits, 256 entries, 0x000000ff mask, 1024 bytes
Rules: 100
Workload: 256 loops, 80 accel
Watchdog: Temperature abort trigger set to 90c
Watchdog: Temperature retain trigger set to 80c
Device #1: Quadro 600, 1023MB, 1200MHz, 2MCU
Device #1: Kernel .../kernels/4318/00000_00.sm 21.32.ptx

Starting attack in stdin mode...

cb7a1ad5035f61b3437c979494f8cd2:351554051050494
Session.Name...: cudaHashcat-plus
Status.....: Cracked
Rules.Type....: File (PREFIX_IMEI), File (theRule2digit)
Input.Mode....: Pipe
Hash.Target...: cb7a1ad5035f61b3437c979494f8cd2
Hash.Type.....: MD5
Time.Started...: Mon Jun 10 15:25:02 2013 (1 sec)
Speed.GPU.#1...: 81007.7s/s
Recovered.....: 1/1 (100.00%) Digests, 1/1 (100.00%) Salts
Progress.....: 53240000
Rejected.....: 0
Memon.GPU.#1...: 99% Util, 53c Temp, 30% Fan

Started: Mon Jun 10 15:25:02 2013
Stopped: Mon Jun 10 15:25:03 2013
```

It's NOT:

- ▶ It takes less than one second to deanonymise on a regular PC if smartphone manufacturer and model are known (which is the case here!)

Sofialys: confirmation of its presence in the App

Below is the listing containing class descriptors of Android App

```
Class descriptor : 'Lnet/hockeyapp/android/UpdateActivityInterface;'  
Class descriptor : 'Lnet/hockeyapp/android/UpdateInfoAdapter$1;'  
Class descriptor : 'Lnet/hockeyapp/android/UpdateInfoAdapter;'  
Class descriptor : 'Lnet/hockeyapp/android/UpdateInfoListener;'  
Class descriptor : 'Lnet/hockeyapp/android/UpdateManager;'  
Class descriptor : 'Lnet/hockeyapp/android/UpdateManagerListener;'  
Class descriptor : 'Lnet/hockeyapp/android/VersionCache;'  
Class descriptor : 'Lcom/adbox/AdBoxLibrary$6;'  
Class descriptor : 'Lcom/adbox/AdBoxLibrary;'  
Class descriptor : 'Lcom/adbox/beans/BanniereDynamique;'  
Class descriptor : 'Lcom/adbox/beans/BanniereExtensible;'  
Class descriptor : 'Lcom/adbox/beans/BanniereRetractable;'  
Class descriptor : 'Lcom/adbox/behavior/DynamicAdBehavior;'  
Class descriptor : 'Lcom/adbox/display/DisplayRetractableBanner;'  
Class descriptor : 'Lcom/adbox/imgthread/ImgException;'  
Class descriptor : 'Lcom/adbox/parsethread/ParseException;'  
Class descriptor : 'Lcom/fabernovel/ratp/AbstractWebMapActivity;'  
Class descriptor : 'Lcom/fabernovel/ratp/AlertingActivity;'  
Class descriptor : 'Lcom/fabernovel/ratp/DetailsTrafic;'  
Class descriptor : 'Lcom/fabernovel/ratp/DetailsTravaux;'  
Class descriptor : 'Lcom/fabernovel/ratp/FDRoute;'  
Class descriptor : 'Lcom/fabernovel/ratp/HorairesResultats;'  
Class descriptor : 'Lcom/fabernovel/ratp/PlansAffichage$StationsOverlay;'  
Class descriptor : 'Lcom/fabernovel/ratp/ProximitePlan$StationsOverlay;'  
Class descriptor : 'Lcom/fabernovel/ratp/Trafic;'  
Class descriptor : 'Lcom/fabernovel/ratp/entity/BusStop;'  
Class descriptor : 'Lcom/fabernovel/ratp/entity/Station;'
```

Responsibility of Google

- ▶ The Android permission system cannot be interpreted as an informed end-user agreement for the collection and use of personal data by third- parties.
- ▶ Android doesn't provide an option for the user to choose the permissions; the user needs to give all the permissions to the App or otherwise, he must just stop using the App.
- ▶ A&A libraries included by the App developer have access to the same set of user PI as the App itself.
- ▶ Behavioral analysis is required.
- ▶ Permission system must be more granular

Outline

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App
- 4 Conclusions**
- 5 References

The answer of RATP (added on July 5th, 2012)

RATP wishes to reply in light

What!!!

They collected WiFi MAC Address,
Device name (Jagdish's iPhone)
among other kinds of info...

items of additional information:

Data exchanges with the

- SDK Sofialys (the advertising SDK) sends information to the Adgoji server, the Fly Targeting system that provides contextual information based on the applications installed on the terminal. Information recovered by SDK is processed for analysis, but does not make it possible under any circumstances to identify the data user.

- **No data collected by Adgoji** concerning users of the RATP app have been used. The Fly Targeting module was under study in Sofialys, which mistakenly implemented it in its SDK in "production" phases. **We are currently removing it from SDK.**

Furthermore, **we confirm that no personal data are used.** In accordance with Apple directives, the UDID stopped being used last year. As for the IMEI: although the ID is already hashed, we are requesting a new revision of SDK.

Why should someone collect
the info they don't use?

- ▶ There is a clear need of better regulations
- ▶ People must understand privacy better

- ▶ We discuss bad practices employed in the world of smartphones (RATP Android and iOS Apps are good illustration)
 - ➊ A&A companies are using not-supposed-to-be ways to collect user PI and tracking mechanisms
 - ➋ They're one step ahead of the OS providers (blocking access to a set of tracking mechanisms lead to shift to some new tracking mechanisms)
- ▶ We discuss the limitations of the privacy control features proposed by Android/iOS Mobile OSs
- ▶ Above all, this is happening without user knowledge.

- 1 Introduction
- 2 RATP iOS App
- 3 RATP Android App
- 4 Conclusions
- 5 References



Apple URL Scheme Reference.

https://developer.apple.com/library/ios/featuredarticles/iPhoneURLScheme_Reference/Introduction/Introduction.html.



MobileSubstrate.

<http://iphonedevwiki.net/index.php/MobileSubstrate>.



On the Effectiveness of Dynamic Taint Analysis for Protecting Against Private Information Leaks on Android-based Devices.

http://www.nicta.com.au/pub?doc=7091&filename=nicta_publication_7091.pdf.



Path uploads your entire iPhone address book to its servers.

<http://mclov.in/2012/02/08/path-uploads-your-entire-address-book-to-their-servers.html>.



Sofialys.

<http://www.sofialys.com/en/>.



Taintdroid.

<http://appanalysis.org>.



Theos.

http://iphonedevwiki.net/index.php/Theos/Getting_Started.



Twitter mobile apps storing address books for 18 months.

http://www.theregister.co.uk/2012/02/15/twitter_stores_address_books/.



WSJ: What They Know - Mobile.

<http://blogs.wsj.com/wtk-mobile/>.