Lalaine: Measuring and Characterizing Non-Compliance of Apple Privacy Labels

Yue Xiao, Zhengyi Li, Yue Qin, Xiaolong Bai, Jiale Guan,
Xiaojing Liao, Luyi Xing

ARTIFACT EVALUATED usenix ASSOCIATION AVAILABLE

ARTIFACT EVALUATED usenix ASSOCIATION FUNCTIONAL

Abstract

Apple's app privacy labels aim to help users easily understand an app's privacy practices. However, misleading labels can trick privacy-aware consumers into data-intensive apps, diminishing the labels' credibility. Lalaine is the first systematic study to evaluate the consistency of data-flow to privacy labels, analyzing 5,102 iOS apps to assess the extent of label non-compliance and its implications.

Contribution

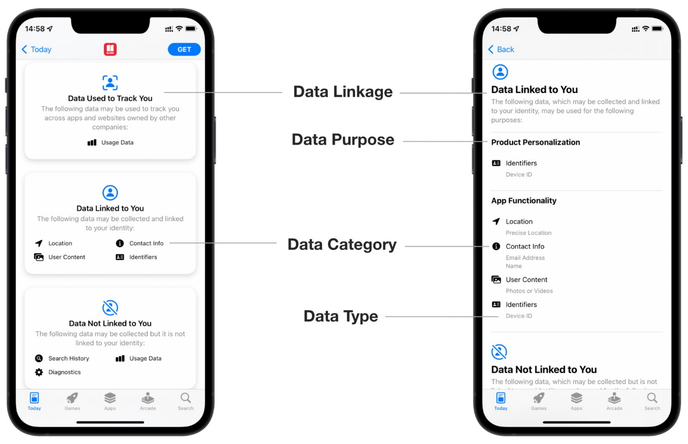
Lalaine - the first large-scale and comprehensive study for privacy label compliance.

- A formalized consistency model
- An end-to-end detection tool
- Root causes analysis & cases study

Apple's Stance on Privacy: A Paradigm Shift –Privacy Label

14 Data Category
Contact Info
Health & Fitness
Financial Info
Identifiers

32 Data Type
Email address
Phone Number
Physical Address
Device ID



4 Data Linkage
Data Used to Track You
Data Linked to You
Data Not Linked to You
Data Not Collected
...

6 Data Purpose
Third-Party Advertising
Analytics
Product Personalization
App Functionality

Consistency Model

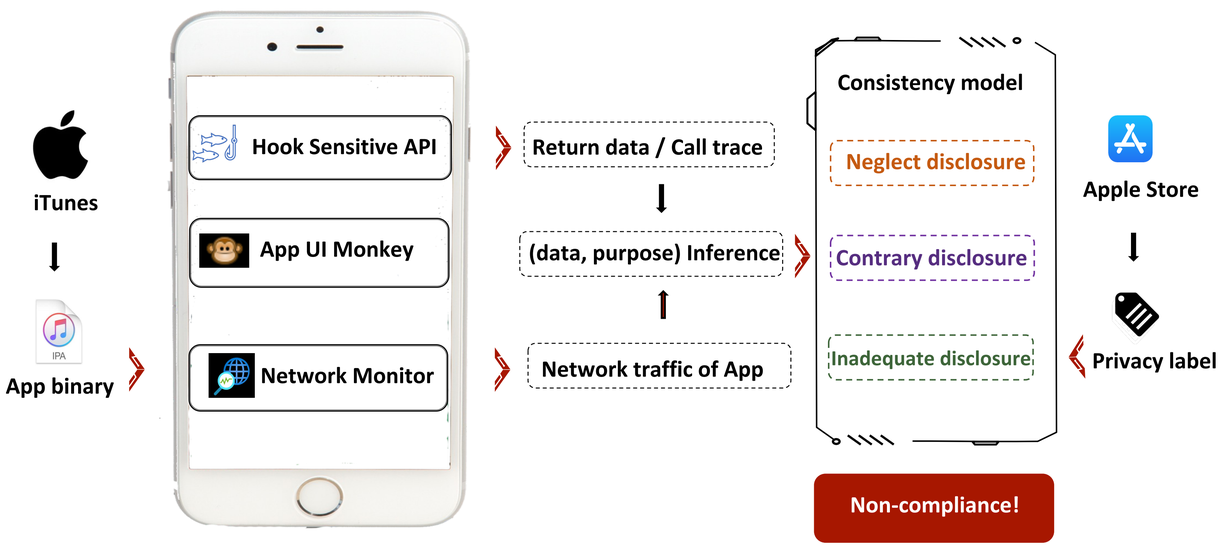
- Data protection principles: data minimization, purpose limitation
- Consistency model: a type of policy language to facilitate automatic compliance check.

- what data is collected from your app and how it is used
- even if you collect the data for reasons other than analytics or advertising, it still needs to be declared

Consistency model

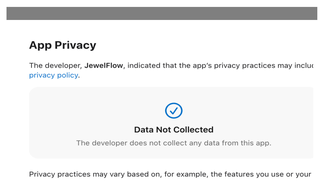
Disclosure representation	$\{s \mid s : (d, q)\}$
Data Flow Representation	$\{f \mid f : (d, q)\}$
Neglect disclosure	$S_f = \emptyset$
Contrary disclosure	$S_f \neq \emptyset \wedge q_f \notin Q_f^S \wedge Q_f^S \nsubseteq Q_f$
Inadequate disclosure	$S_f \neq \emptyset \wedge q_f \notin Q_f^S \wedge Q_f^S \subset Q_f$

Design & Implementation



Real-world Non-complaint apps

Non-compliant App 1: Kanchan Jewellers



Precise Location


```
{
 "app_id" : "7ab536ba-cfd9-4c3e-90a0-7f2e1d43b446",
 "loc_bg" : false,
 "loc_acc_vert" : 10,
 "lat" : 39.174660633583997,
 "net_type" : 0,
 "loc_acc" : 65,
 "long" : -86.491791630912374
}
```


Device ID, Time zone, Device metadata


```
{
 "app_id" : "7ab536ba-cfd9-4c3e-90a0-7f2e1d43b446",
 "net_type" : 0,
 "as_id" : "0F41A05D-B6F8-49FD-AAFF-3D4711371B81",
 "ios_bundle" : "com.triologic.jewelflowpro.kanchanProduction",
 "device_type" : 0,
 "rooted" : true,
 "sdk_type" : "native",
 "sdk" : "021503",
 "device_os" : "13.7",
 "language" : "en-US",
 "game_version" : "3",
 "timezone" : -18000,
 "ad_id" : "5065E8FB-964B-4178-9BCD-F6DCB28082EB",
 "notification_types" : 31,
 "device_model" : "iPhone10,4",
 "identifier" : "5b2e10e020ddad343c..."
}
```


User activity


```
{
 "net_type" : 0,
 "active_time" : 81.500266075134277,
 "state" : "ping",
 "app_id" : "7ab536ba-cfd9-4c3e-90a0-7f2e1d43b446",
 "type" : 1,
 "device_type" : 0
}
```


Email, Full name, Phone number


```
GET /online_api/2.0/UserRegistration/generate_verification_keys?company_code=S0FOQ0hBTg%3D%3D&country_code=%2B1&email_id=hdiaosnd%40gamil.com&full_name=Yue&mobile_no=8123250806&string_random_internal=3857599184 HTTP/1.1
```


Non-compliant App 2: Atlanta News from 11Alive

Inconsistency Type	Endpoint	Data Collection (example)	Actual Purposes	Disclosed Purposes
Neglect Disclosure	https://www.google-analytics.com/collect?	user_id=83360165816509286981 ios_jb=1, sz=300x250	Analytics	-
Contrary Disclosure	https://pubads.g.doubleclick.net/gampad/ads?	device_id=5B30BC06-9017-4FA0-8A77-3FB3FFBE3D7D	Third-Party Advertising	App Functionality
Inadequate Disclosure	https://aax-us-east.amazon-adsystem.com/e/msdk/ads?	*"geoloc":"39.154663, -86.492607, 65.000000, 1"*	Third-Party Advertising	App Functionality
	https://api.tegnadigital.com/mobile/configuration-ro/updateUserLocation?	"lastKnownCoordinates": {"latitude":39.136590224210444, "longitude":-86.48033368434717}	App Functionality	

Table 1: Three inconsistencies in app *Atlanta News from 11Alive*

Privacy label

App Functionality
- Location — Precise Location
- Contact Info — Email Address, Name
- User Content — Photos or Videos
- Identifiers — Device ID

Result

Running on 5,102 iOS apps, Lalaine detects 3,423 privacy label non-compliance.

Root cause analysis

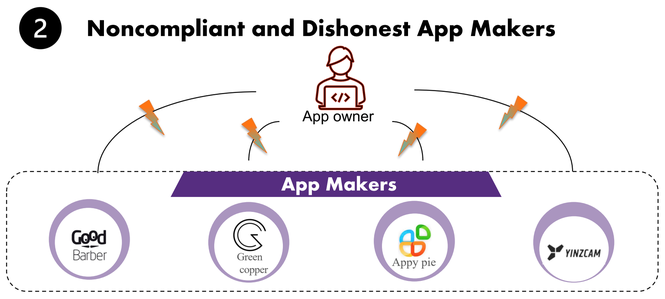
1 Misleading Third-Party Disclosure guidelines

Browsing history	No	Don't collect
Search history	No	Don't collect
Identifiers • User ID • Device ID (IDFA)	Optional	• User ID: App developers may set and transmit user ID (CUID) • Device ID: The Advertising ID (IDFA) is collected only when accessible.
Purchases • Purchase history	Optional	App developers can choose to configure certain in-app purchase events to measure.
Usage data • Product interaction • Advertising data • Other usage data	• Yes • Optional • No	• Product interactions: App launches are measured, as well as any other user interaction configured by the app developer. • Advertising data: is received if the app developer uses the AppsFlyer ad revenue module.

Caller API	System API
[FlurryLocationInfoSource onqueue_dataProviderDidUpdate]	[CLLocation - coordinate]

Network traffic	Endpoint
{"fl.altitude.value":0.5198214054107666,"fl.location.permission.status":true,"fl.speed.value":-1,"fl.longitude.value":-122.25865937775535,"fl.bearing.value":-1,"fl.latitude.value":37.552978515625,"fl.horizontal.accuracy.value":65,"fl.vertical.accuracy.value":10,"fl.report.location.enabled":true,"fl.time.epoch.value":1653769294876}	https://data.flurry.com/v1/flr.do

2 Noncompliant and Dishonest App Makers



Company	com.goodbarber	com.greencopper	com.appypie	com.yinzcam
Collected data	User ID, Precise/Coarse Location, Other Diagnostic Data	Sensitive Info	Device ID, User ID, Other Diagnostic Data,	Advertising Data
Endpoint	api.ww-api.com	user-state.greencopper.com	api.appexecutable.com	ads-min-mls.yinzcam.com
#Non-compliant apps	85	18	15	5

Take aways

Pervasive non-compliance app

- The privacy label non-compliance in iOS apps are prevalent, with a serious impact on credible and transparent disclosure of app privacy practices.
- The root causes of privacy label non-compliance are diverse.
- We are reporting all findings (non-compliant privacy labels) to Apple.

An end-to-end detection tool

We released the end-to-end tool (including UI automation, data purpose inference, consistency check) to automatically assess the "flow-to-label" inconsistency

Recommendations for Stakeholders

- Apple: A comprehensive ontology of sensitive data items/sensitive API
- SDK Vendors: improve data transparency and guidance accuracy for privacy labels.
- App Developer: scrutinize SDKs to understand their data practices