スマートフォンからの利用者情報の送信
～情報収集の実態調査～

KDDI研究所
竹森 敬祐

なぜ利用者情報の送信を危惧しているのか？
仕様がオープンな Android™フォンを例に解説します
が、送信はAndroid™フォンに限った話ではありません。

1: Android™のセキュリティを学ぶ

2: 実態調査

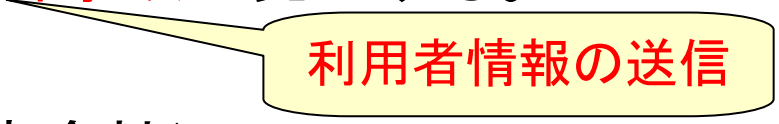
注）KDDI研究所開発ツールを用いた調査であり、抜けや誤りについてはご容赦ください。

KDDI R&D LABS

はじめに

- スマートフォンとプライバシ
  - 個人との結びつきが強く、利用者を映し出すPCである。

- Android™OSの思想（利便性）
  - アプリが利用できる機能や情報が豊富で、便利なアプリを実現。
    ⇒ 問題を含むアプリや悪意のアプリが開発される。
  - Android™フォンは、携帯電話の識別子を持ったPCである。
    ⇒ PCと同様＋αのセキュリティ事故が発生する。

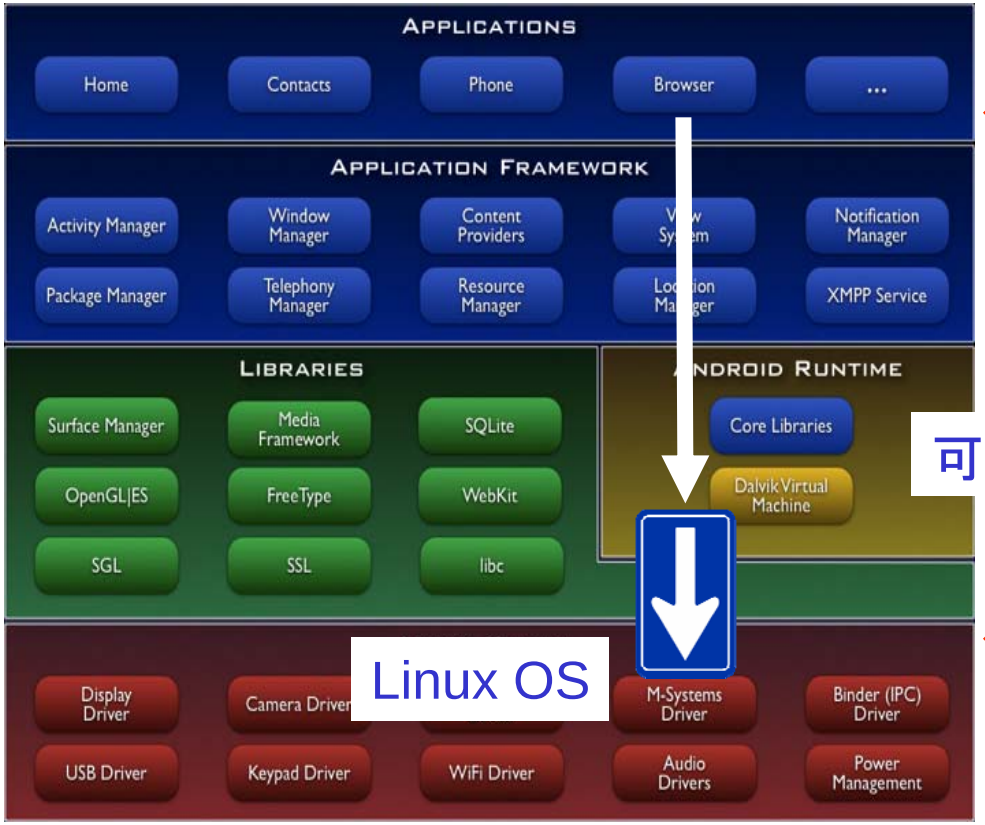


- Android™OSのセキュリティ（安全性）
  - PCのような自動的なウイルス感染は殆どない（Android™にウイルス無し）。
  - インストール時にアプリの権限（パーミッション）を確認・承認する。
    ⇒ アプリの良性／悪性の判断は難しい。
  - サンドボックスで隔離されて実行される。
    ⇒ アプリの挙動をモニタすることは難しい。

KDDI R&D LABS

Android™OSのセキュリティ機構 ～サンドボックス～

- Linux OS + サンドボックス
  - Linux上に、パーミッション可変型のサンドボックスを構築したOS。
  - アプリが利用する機能・情報をユーザが承認する。



- 情報アクセスへの機能が豊富
  ⇒アプリ開発者による誤適用

可変型（ユーザ承認型）サンドボックス

- パーミッションの作用はアプリ開発者次第
  ⇒ ユーザの誤判断

KDDI R&D LABS

Android™OSから求めるユーザ承認

■ パーミッション機構

◆ Android™OSから、アプリが利用する機能や情報を表示して、ユーザ承認を求めるインストール機構となっている（Android Market™はFunction表示でユーザ許諾を得る）。

⇒ 機能や情報を利用する目的が記されていない。

⇒ 機能や情報単位で申請であり、総合的な作用や悪意の判断は難しい 。

★ Android™のパーミッションで説明済みだが、アプリ側でのフォローも望まれる。

スマートフォンからの利用者情報の送信
～情報収集の実態調査～

KDDI研究所
竹森 敬祐

なぜ利用者情報の送信を危惧しているのか？
仕様がオープンな Android™フォンを例に解説します
が、送信はAndroid™フォンに限った話ではありません。

1: Android™のセキュリティを学ぶ

2: 実態調査

注）KDDI研究所開発ツールを用いた調査であり、抜けや誤りについてはご容赦ください。

KDDI R&D LABS

ところで、マルウェアの出現状況

- PC・モバイルの総計（2011年1月～2011年10月観測）
  - のべ1,300,000種類、約4,300種類／日の検体を観測。

- Androidの統計（2011年1月～6月観測）
  - のべ200種類、約1.1種類／日の検体を観測。

マルウェア出現数　PC：Android™フォン ＝ 4000：1

（Android™フォンのマルウェア感染は殆ど無く、PCよりもかなり安全）

勝手な情報送信の主原因は、アプリ開発者による情報収集モジュールの誤用です。

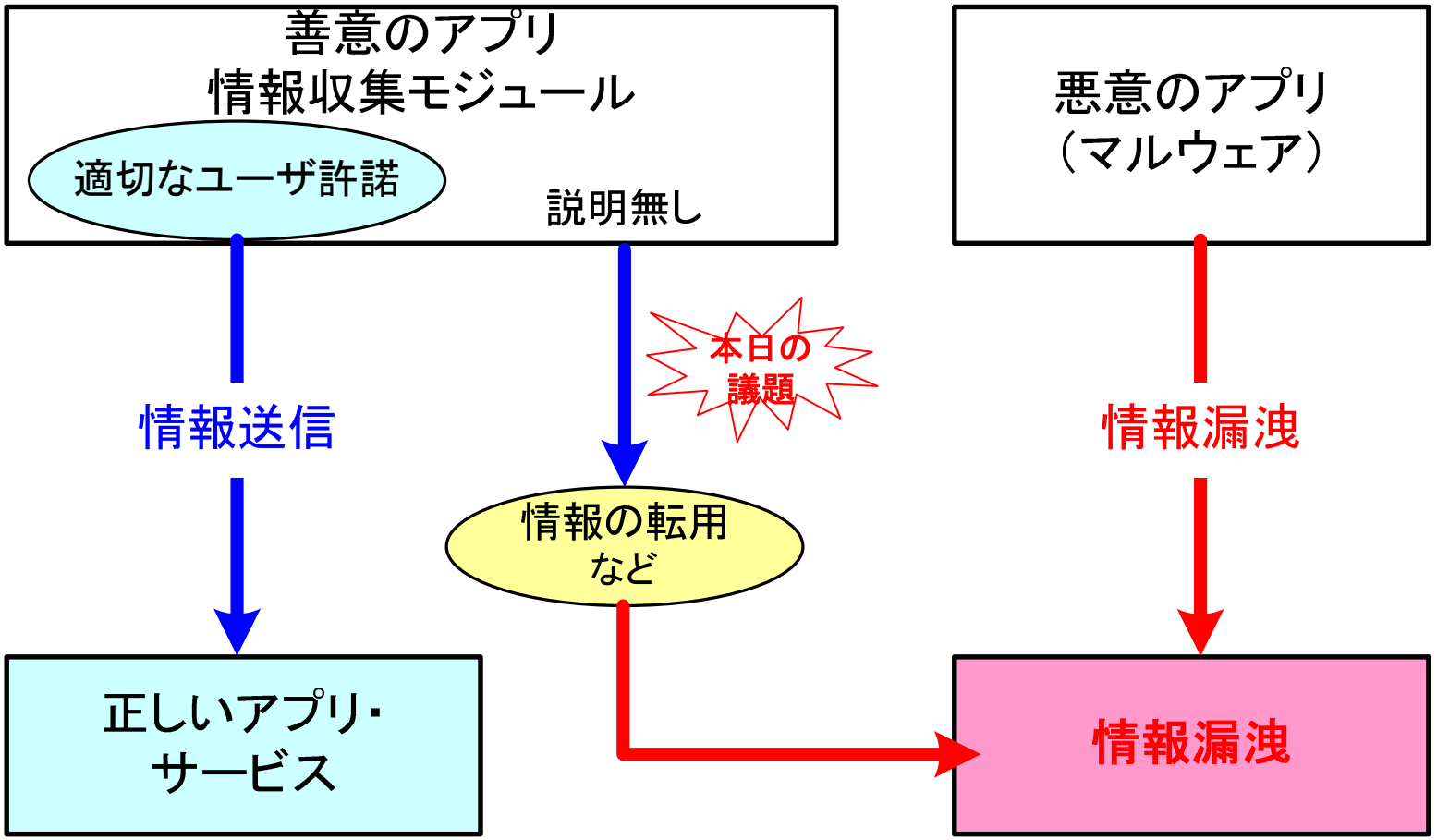
マルウェア対策ソフトの検知対象外

統計情報の提供元
株式会社カスペルスキー

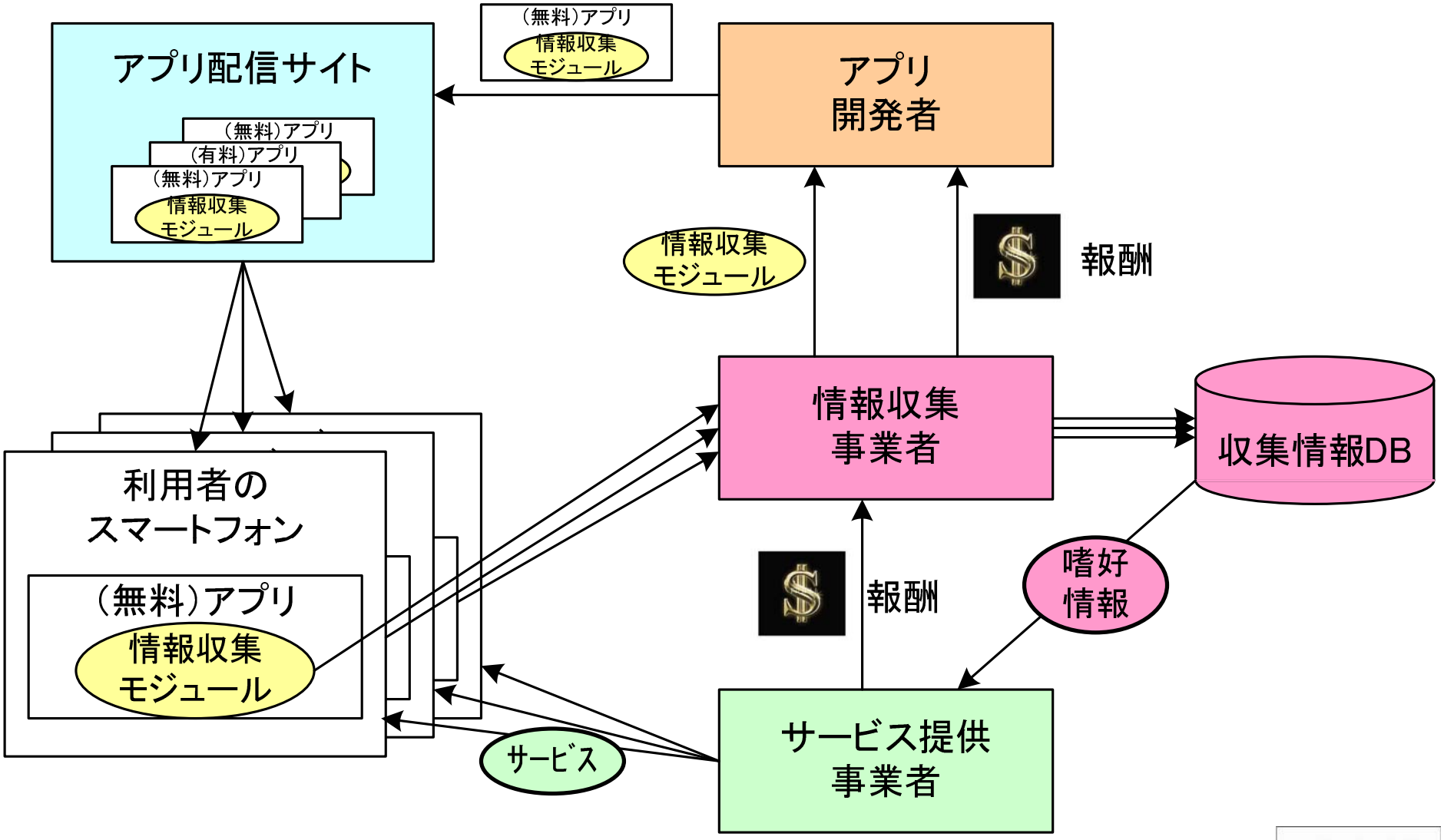
本日の議題の整理

■ 利用者情報の送信例

◆ 利用者への説明や許諾を伴わない情報の収集や転用、マルウェアによる漏洩。



情報収集ネットワークの一例



Android™フォンにおける利用者情報

表1. Android™フォンから送信できる利用者を特定する情報および各種識別子（ID）

種別	詳細
利用者を特定する情報	氏名、アドレス帳で管理される情報、メールアドレス
個体を識別する情報（ID）	OSが生成するID（Android ID：0x16桁）、端末ID（IMEI：d14桁）、加入者ID（IMSI：d14桁）、SIMシリアルID（ICCID：d19桁）、電話番号（d11桁）、認証チケット（AuthToken）、アプリケーションが独自に発行するID、MACアドレス、OSやサービスへのログインアカウント、IPアドレス＊1など
ハッシュ値	IDのハッシュ値＊2 ＊3

＊1) 通信キャリアはIPアドレスで端末（利用者）を識別できうる。
＊2) 桁数の少ないIDのハッシュ値は元のIDを有限時間で探索できる。
＊3) 誰もが同じ値を算出できるIDからのハッシュ値は共通IDと考えることもできる。

表2. Android™フォンから送信できるプライバシ情報

種別	詳細
利用履歴	位置情報、通話の内容・履歴、メールの内容・履歴、Webのブックマーク・閲覧履歴など
アプリケーション	アプリケーションの一覧・利用履歴、アプリケーションの管理データ＊4など
システム	スマートフォンのシステムログ

＊4) アプリケーションの管理データは、業務資料や写真など、アプリケーションの仕様によって様々である。

KDDI R&D LABS

利用者情報アクセスに関わるAndroid™パーミッション

■ Android Market™の無料アプリ（14カテゴリ×70個＝980個）のパーミッション（2011/8アプリ）

表3. 特に注意の必要なPermission

利用率(%)	種別	取得できる情報
57.9	READ_PHONE_STATE	端末ID（IMEI）、加入者ID（IMSI）、SIMシリアルID（ICCID）、電話番号、通話相手の電話番号
28.4	ACCESS_COARSE_LOCATION	基地局・WiFiを使った位置情報
26.4	ACCESS_FINE_LOCATION	GPSを使った位置情報
10.6	READ_CONTACTS	アドレス帳（氏名、電話番号、メールアドレス、住所など）
9.1	GET_TASKS	実行されたアプリ名
8.3	READ_LOGS	実行されたアプリ名、通話履歴、Webアクセス履歴など
8.2	GET_ACCOUNTS	Googleアカウント（Gmailアドレス）
3.7	USE_CREDENTIALS	Googleアカウントの認証結果（AuthToken）
3.5	READ_SMS	SMSメール（Cメール）

注）種別に記されるパーミッション名には、”android.permission.”が前に付与されます。表サイズ制限のため省略していることに注意ください。

利用者情報アクセスに関わるAndroid™パーミッション

■ Android Market™の無料アプリ（14カテゴリ×70個＝980個）のパーミッション（2011/8アプリ）

表4. 注意すべきではあるが考察を要するPermission

利用率(%)	種別	取得できる情報
16.3	ACCESS_WIFI_STATE	Wi-Fiアクセスポイント情報
10.1	CAMERA	カメラ撮影
4.8	RECORD_AUDIO	録音
2.8	READ_CALENDAR	Googleカレンダ情報
2.4	READ_HISTORY_BOOKMARKS	Webアクセス履歴・ブックマーク
1.5	AUTHENTICATE_ACCOUNTS	Googleアカウントのパスワード
0.5	READ_OWNER_DATA	利用者情報
0.4	ACCOUNT_MANAGER	Googleアカウント情報
0.0	READ_FRAME_BUFFER*	スクリーンショット
0.0	READ_INPUT_STATE*	キー入力

注）種別に記されるパーミッション名には、”android.permission.”が前に付与されます。表サイズの制約のため省略している。
* 一般権限のアプリから利用することはできません。

Android™パーミッションの不要な利用者情報

表5. パーミッションの許諾なくアクセスできる情報

種別	詳細
Android ID	OSが初回起動時に生成する16桁の乱数＝端末IDとみなせる
アプリ名	インストールされているアプリー覧
SDカード	SDカード上で管理される情報（アプリのデータなど）

情報収集モジュールの実態

■ ターゲット広告

◆ アプリ内の広告をユーザがクリックすることで、開発者に報酬が入る。

⇒ READ_PHONE_STATE: Android ID、電話番号などからユーザを識別。

⇒ ACCESS_COARSE(FINE)_LOCATION：場所に応じた広告を表示。

■ 望ましい姿

◆ Android™の安全機構（パーミッション）でユーザから承諾は得ているものの、情報収集モジュールを組み込んだアプリ開発者はユーザに対して収集する情報、利用目的や範囲などをアプリの中で説明／許諾を得た方が親切。

検索アプリ

■ 情報収集モジュールの含有実態

◆ Android Market™の14カテゴリ×70個＝980個の無料アプリを対象に含有する情報収集モジュールを調査。

	含有数	含有率
アプリ総計	558/980	56.9%
情報収集モジュール総計	1065/558	1.91個

KDDI R&D LABS

情報収集モジュールの統計（2011/8 調査）

■ KDDI研は、Android Marketから14カテゴリ×70件＝980のアプリを取得・調査した。

注）組み込みモジュールの統計であり、送信情報の統計ではない。

表6. 980アプリから抽出された情報収集モジュール

情報収集モジュール一覧	外部送信を確認した情報	対象アプリ 件数	980アプリ 利用比率
com.	AndroidId, 国名, 端末名	269	27.45%
com. ds	AndroidId, AndroidId(ハッシュ値), IMEI, 国名, 端末名	212	21.63%
com.	AndroidId, IMEI, 位置, 端末名	86	8.78%
com. pps/analytics	国名, 端末名	83	8.47%
com.	AndroidId(ハッシュ値), 国名	58	5.92%
com.	–	58	5.92%
com. ndroid	IMEI, 国名, 端末名	44	4.49%
net/	–	40	4.08%
com.	–	39	3.98%
com. sdk	AndroidId, IMEI, 国名, 端末名	38	3.88%
jp/c r	AndroidId(ハッシュ値), 国名, 端末名	24	2.45%
com. roid	電話番号, AndroidId, 端末名	23	2.35%
com.	AndroidId(ハッシュ値), 端末名	16	1.63%
com. k	AndroidId, 国名, 端末名	16	1.63%
com.	IMEI, 国名, 端末名	12	1.22%
com.	–	12	1.22%
com. /ads	AndroidId, 国名,端末名	10	1.02%
com.	AndroidId, 位置, 端末名	8	0.82%
com. d/sdk	IMEI	7	0.71%
com. ndroid	電話番号, AndroidID	6	0.61%
com.	–	3	0.31%
com.	–	1	0.10%

送信情報の統計（2011年8月アプリ、2011/12-2012/1評価）

■ 980アプリのうち400アプリについて5分間の挙動解析

表7. 送信を確認した情報

	件（率）/400	送信情報
ID	50件(12.5%)	Android ID
	57件（14.3%）	端末ID（IMEI）
	7件（1.8%）	加入者ID（IMSI）
	0件（0.0%）	SIMシリアルID（ICCID）
	7件（1.8%）	Googleアカウント（Gmailアドレス）
	87件（21.8%）	Android IDのMD5ハッシュ値
	4件（1.0%）	IMEIのMD5ハッシュ値
	4件（1.0%）	電話番号
プライバシ	32件（8.0%）	位置（緯度・経度）
	3件（0.8%）	アプリ一覧

注）研究開発ツールの実行時ログから抽出したものであり、抜けや誤りがあることをご容赦ください。

KDDI R&D LABS

送信情報の統計調査（2011年8月アプリ、2011/12-2012/1評価）

■ 何らかの情報を送信するアプリ
- ◆ 表7のいずれか1つ以上の情報を送信しているアプリ：181/400件（45.3%）
  - ⇒ 14/400件（3.5%）が「説明」あり　うち適切な説明が 8/400件（2.0%)
  - ⇒ 10/400件（2.5%）が「許諾」あり　うち適切な許諾が 9/400件(2.3%)
  - ★ 164/400件（41.0%）が（説明｜許諾）なし、もしくは適切な（説明｜許諾）なし

■ ID+プライバシに関わる情報を送信するアプリ
- ◆ 表7の青＋黄の組合せ情報を送信しているアプリ：31/400件（7.8%）
  - ⇒ 5/400件（1.3%）が「説明」あり　うち適切な説明が 4/400件（1.0%）
  - ⇒ 2/400件（0.5%）が「許諾」あり　うち適切な許諾が 2/400件（0.5%）
  - ★ 25/400件（6.3%）が（説明｜許諾）なし、もしくは適切な（説明｜許諾）なし

Android™の安全機構（パーミッション）で利用者へ説明済みであり、上記のアプリが直ぐに悪い訳ではない。説明や許諾をアプリ内に追記した方が良いものが6.3%あるという趣旨である。

注1）適切な説明や許諾とは、具体的な送信情報＋送信先＋利用目的が適切なタイミングであるもの。
注2）56.9%のアプリに情報収集モジュールがあり、45.3%のアプリが何らかの情報を送信。
　11.6%の差異：秘匿して送信しているもの、国をみて停止するもの、アプリの収集日と評価日のズレ？

KDDI R&D LABS

情報収集モジュールの含まれるアプリ例「許諾なし」

- パーミッション
  - ACCESS_COARSE_LOCATION、READ_PHONE_STATE

- 問題点
  - アプリ内での利用者許諾なし（中国の情報収集モジュール）
  - 送信：端末識別ID（IMEI）＋位置情報

GET /kuAD_V2/InfoReceive.php?cmd=REG&apid=00000000e&ver=04&imei=354957031150819 HTTP/1.1¥r¥naccept: */*¥r¥ncontentType: charset=utf-8¥r¥nUser-Agent: Dalvik/1.4.0 (Linux; U; Android 2.3.4; Nexus One Build/GRJ22)¥r¥nHost: [redacted].com¥r

GET /kuAD_V2/InfoReceive.php?cmd=LBS&did=000000taEh&lat=35.87912053333333&lon=139.51742570000002&acc=66.0 HTTP/1.1¥r¥nUser-Agent: Dalvik/1.4.0 (Linux; U; Android 2.3.4; Nexus One Build/GRJ22)¥r¥nHost: [redacted].com¥r¥n

新聞早読みアプリ

KDDI R&D LABS

情報収集モジュールの含まれるアプリ例「不適切な許諾」

■ パーミッションと情報収集モジュール

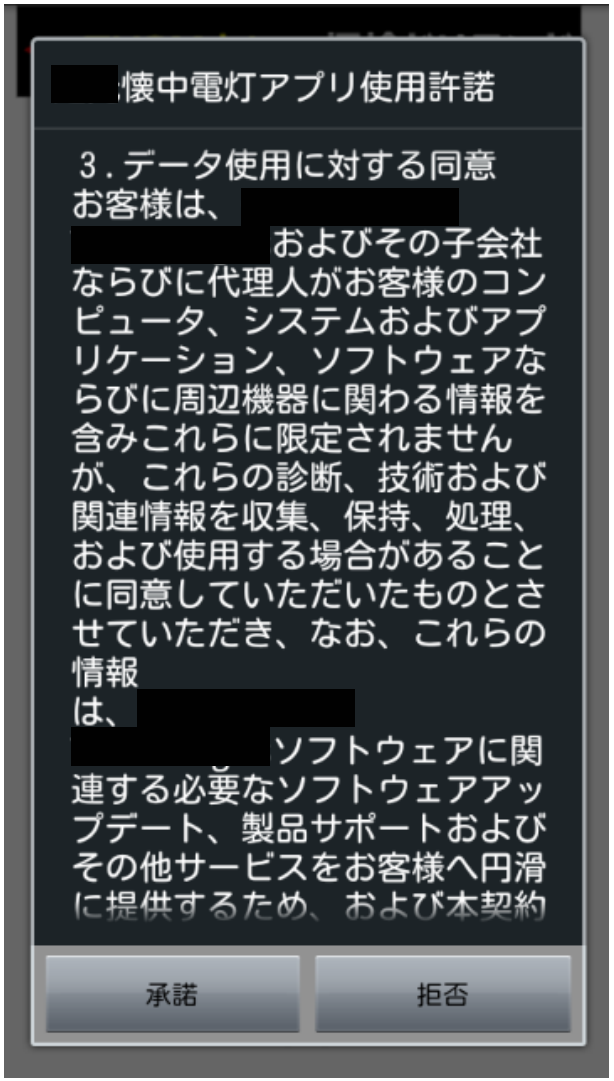
- WAKE_LOCKCAMERA、FLASHLIGHT、STATUS_BAR、WAKE_LOCK、ACCESS_COARSE_LOCATION 、ACCESS_FINE_LOCATION、ACCESS_NETWORK_STATE、INTERNET、READ_PHONE_STATE、WRITE_EXTERNAL_STORAGE、ACCESS_WIFI_STATE、他独自を12個
- 9種類の情報収集モジュールを内包する。

■ 問題点

- 許諾前に情報送信

  「承諾」クリック前にアプリが実行される。
- 説明が不適切

  収集情報の説明が曖昧、収集者と目的が違う。
- 送信： Android ID, 端末ID(IMEI)



懐中電灯アプリ

情報収集モジュールの含まれるアプリ例「不適切な説明」

■ パーミッションと情報収集モジュール

◆ INTERNET、ACCESS_FINE_LOCATION、ACCESS_WIFI_STATE、READ_CONTACTS、VIBRATE、WRITE_EXTERNAL_STORAGE、WAKE_LOCK、GET_TASKS、READ_PHONE_STATE、ACCESS_NETWORK_STATE 、他を3個

■ 問題点

◆ 説明が不適切

自社アプリが扱う収集情報に関する説明のみ。

外部の情報収集モジュールが扱う情報の説明なし。

◆ 送信：端末ID（IMEI）＋位置情報

```
GET /post/config?p=android&a=null&m=2.3.0&v=2.0.
1&d=a0000008cf97fd&dm=IS06&dv=2.2.1&hwdm=j
masai&g=wifi&ll=35.8789315333333 36%2C139.517
35818333333&mcc=440&mnc=127&new=true HTTP
/1.1¥r¥nCookie: ¥r¥nUser-Agent: null¥r¥nHost: data.
[redacted].com¥r¥nConnection: Keep-Alive¥r¥n¥r¥n
```

ゲームアプリ

KDDI R&D LABS

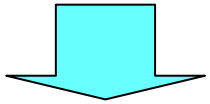
課題と実態の乖離

ニコニコ生放送での質問
実際に脅迫・金銭被害はあるのか？

本件の実被害はゼロ

■ KDDI研究所の取り組み
- ◆ 情報収集モジュール組込みアプリおよびマルウェアをクリック。
- ◆ 既に2年が経過するが、脅迫の被害に遭ってない。

⇒ Android™アプリの勝手な情報送信による被害の実態を掴めない。



★ Android™アプリを利用している中で、個人情報の漏洩と思われる脅迫などの具体的な被害に遭われた方は、いらっしゃいますか？

⇒ 何のアプリですか？

⇒ どのような脅迫でしたか？

⇒ 言語は何でしたか（英語、中国語、日本語etc）？

注）問題の切り分けのため、具体的な事象をお知らせください。

結果）放送中に1件も事故に関する報告は寄せられなかった。

KDDI R&D LABS

スマートフォン特有の問題

- ID送信の考察
  - IDのみが送信された場合でも、何のアプリから送信されたか知りえる。
- 固有の情報
  - 情報収集のAPIが決まっており、情報取得が容易。
    （PCのWebブラウザはサンドボックスが強固 ⇔ Android™フォンでは利用者の承認でサンドボックスに穴が開く）
  - PCには無かった電話番号、アドレス帳を持つ。位置の特定も容易。
  - 利用者側で破棄できない固定値（世界基準のID）がある。
- Always On／常に携帯
  - ネットに常時繋がるため、サーバ連携型のサービスが多い。
  - 個人との結びつきが強く、ターゲティングサービスを提供しやすい。
- 利用者層
  - PC利用の経験のない利用者層まで、スマートフォンが普及する。

根本原因

■ アプリ開発者の収入源
- ◆ アプリ開発者の裾野が、全世界的に、個人層まで拡大している。

⇒ 情報管理に対するリテラシの低いアプリ開発者も混じっている。

⇒ 情報収集モジュールの特性を理解しないままアプリに組み込む。

■ Android™OSのパーミッション機構
- ◆ 機能毎の通知が限界であり、情報収集の意図が伝わりきらない。

⇒ 利用者情報の収集について、十分な許諾を得られたとアプリ開発者が勘違いしている。

■ 世界のマーケットからアプリが流入
- ◆ 殆どのアプリは、日本の法規制を受けない。

⇒ ユーザ許諾のあり方を日本だけで議論しても解決しない。

付録：KDDI（au one Market）の検査ツール：情報収集モジュール

付録：情報収集モジュールのメリット・デメリット

	種別	メリットの例	デメリットの例
ユーザ	広告	・アプリの無料化/低価格化 ・個人にあった広告/推薦情報の表示	・外部に送信・蓄積された利用者情報が悪用される恐れ
	サービス	・子供の見守りサービス等の位置追跡 ・アルバム/カレンダー等のネット管理	
アプリ開発者 サービス提供者	広告	・広告クリックによる対価の受取り ・嗜好把握による適切な広告提供	・説明不足で信頼失墜の恐れ ・情報管理における事故
	サービス	・見守りサービスなどの開発/提供 ・統計分析による戦略/立案	