5W-02

動的解析によるAndroidアプリが取得しているプライバシー情報の調査
Study on Privacy Information Collected in Android Apps with Dynamic Program Analysis

程　斌† 角田　裕太‡ 細谷　竜平‡ 森　達哉§ 齋藤　孝道†
明治大学† 明治大学大学院‡ 早稲田大学§

1 はじめに

Androidアプリでは，APIを利用して端末利用者に関する情報および端末を特定できる情報を収集し，外部へ送信することがある．それらによって利用者の個人情報が意図せずにサードパーティに送信される可能性がある．そこで，本論文ではAndroidマーケットでリリースされている人気アプリに対してJavaのデバッグインターフェースを利用した動的解析を行い，Androidアプリが取得しているプライバシー情報について調査を行った．調査の結果，調査対象とした34個のアプリ全てで，Android IDとUUIDが取得されていた．そのうちAndroid IDをサードパーティに送信しているアプリが23個，UUIDをサードパーティに送信しているアプリが25個であった．その他，IMEIやWiFiの対応情報やGPS情報などの情報を取得，送信しているアプリも多数確認した．

2 関連研究

先行研究において，Androidアプリの情報取得に関する調査が行われている．

細谷ら[1]は，コード解析を利用してアプリケーションが取得している情報をAPIレベルで分析した．細谷らは国内でリリースされている1,704個のアプリに対し，コード解析を行った．その結果，40.08%は端末のシリアル番号を，71.06%は位置情報を，39.26%はインストール済みアプリケーション名を取得していることを示した．

福田ら[2]はJDWPと呼ばれるAndroidアプリにおけるデバッグ情報のやり取りを行うプロトコルを用いて，Androidアプリに組み込まれている外部モジュールが外部へ送信しているグローバルIDを調査した．その結果，国内でリリースされているAndroidアプリの11%がIMEIやIMSIといった，Androidアプリにおいて非推奨とされているグローバルIDを送信していたことを示した．

本研究では，動的解析を用いてグローバルIDに加え，利用者が取得されることに嫌悪感を抱く情報のアプリ内での取得状況とサードパーティへの送信状況の調査を行う．

本論文では，Androidマーケットでリリースされているアプリに対し動的解析を行った．この解析の目的は，リリースされているAndroidアプリが利用者の情報を取得し，それを外部へ送信している事実を示し，判断の助けとすることにある．解析の対象としたアプリおよびその解析結果は研究以外の目的で使用されておらず，研究室にて厳重に管理されている．

3 関連知識

3.1 プライバシー情報

本論文で調査対象とするプライバシー情報については端末，ネットワークおよび個人の特定の可能性がある情報に着目する．各アプリにおいて，以下の3種類の情報に対する取得状況を調査した．

端末を特定できる情報

Android ID：Android端末を識別するための識別子である，ファクトリーリセットによって変更できる．

UUID：アプリを一意に識別するために用いられるグローバルIDである，アプリの再インストールによって変更できる．

Ad ID：広告に用いられるグローバルIDである，利用者によってリセットすることができる．

IMEI：携帯電話に紐づく一意な識別子である，基本的に変更できない．

ネットワークに関する情報

WiFiの対応情報：端末がWiFi機能に対応しているかどうかの情報である．

SSID：端末が接続しているWiFiアクセスポイントの名前である．

位置情報

GPS情報：端末のGPSセンサーが取得した位置情報である．

Study on Privacy Information Collected in Android Apps with Dynamic Program Analysis
†Bin CHENG ‡Yuta TSUNODA ‡Ryohei HOSOYA §Tatsuya MORI †Takamichi SAITO
†Meiji University
‡Graduate School of Meiji University
§Waseda University

3.2 Android Debug Bridge

Android Debug Bridge（ADB）は Android 端末とデバッグ情報のやり取りを行うツールである．デバッガはアプリのデバック時に，ADB の通信機能である Java Debug Wire Protocol（JDWP）を介して，デバッグコマンドの通信やアプリのデータ交換などを行う．これにより，デバッガは実行中のアプリのスタックフレーム情報やインスタンスフィールド情報やメソッドの実行情報などを取得することができる．

4 調査方法

本論文では，Google Play ストア [3] でリリースされている各カテゴリのトップアプリに対して動的解析を行った．なお，調査対象としたアプリ内の機能については可能な限り全てにアクセスを行なった．

4.1 調査対象

本論文では，Google Play ストアの各カテゴリ別ダウンロード数ランキング（2017 年 8 月時点）1 位の 34 個のアプリを調査対象とした．

4.2 API 呼び出しの観測

本論文では，Android Studio のデバッガ機能を用いて，Android アプリ実行時の API 呼び出しを観測した．今回，観測対象としたのは 3.1 節に示した情報を取得する API と通信を行う API である．アプリ内でこれらの API 呼び出しを観測することで，プライバシー情報が端末内で取得されているか，またサードパーティに送信されているかを判定することができる．具体的には，ブレークポイントで API を呼び出した時点のスタックフレームを調査し，返り値やローカル変数などを分析することで判定を行う．

5 調査結果

表 1 にプライバシー情報を取得およびサードパーティへ送信している Android アプリの数を示す．

プライバシー情報	取得アプリ数	送信アプリ数
Android ID	34 (100.0%)	23 (67.6%)
UUID	34 (100.0%)	25 (73.5%)
Ad ID	13 (38.2%)	8 (23.5%)
IMEI	11 (32.4%)	0 (0.0%)
WiFi の対応情報	25 (73.5%)	3 (8.8%)
SSID	2 (5.9%)	0 (0.0%)
GPS 情報	26 (76.5%)	0 (0.0%)

表 1: プライバシー情報の取得と送信状況

表 1 より，Android ID や UUID などのグローバル ID が 34 個の人気アプリすべてで取得されていることがわかった．また，これらの情報はそれぞれ 23 個 (約 68%)，25 個 (約 74%) のアプリでサードパーティに送信されていることがわかった．そのほか，SSID や Wifi の対応状況，GPS 情報などのネットワークに関する情報を取得しているアプリはそれぞれ 2 個 (約 6%)，25 個 (約 74%)，26 個 (約 77%) であった．Wifi の対応情報のみサードパーティへの送信が 3 個 (約 9%) あったがそれ以外のアプリではサードパーティへの送信は確認できなかった．

6 まとめ

本論文では 34 種類の人気 Android アプリに対する動的解析を行い，アプリによるプライバシー情報の取得状況とサードパーティへの送信状況を調査した．

調査の結果，Android ID や UUID などの利用者を一意に特定できる可能性のある識別子がすべての人気アプリで取得されていることがわかった．また，これらの情報は，Android ID は 23 個 (約 68%)，UUID は 25 個 (約 74%) のアプリでサードパーティにそれぞれ送信されていることがわかった．また SSID や Wifi の対応状況，GPS 情報などのネットワークに関する情報を取得しているアプリはそれぞれ 2 個 (約 6%)，25 個 (約 74%)，26 個 (約 77%) 存在した．Wifi の対応情報のみサードパーティへの送信が 3 個 (約 9%) あったがそれ以外のアプリではサードパーティへの送信は確認できなかった．今回サードパーティへの送信が確認できた情報の中で Android ID は端末の特定が可能であり，かつ利用者による変更が難しいので端末の追跡に利用される可能性がある．利用者が容易に変更することができない端末識別子の取得は，プライバシー保護の観点からは，問題であると思われる．

参考文献

[1] 細谷 竜平，角田 裕太，森 達哉，齋藤 孝道，モバイルアプリケーションが取得しているプライバシー情報の調査, コンピュータセキュリティシンポジウム 2017(CSS2017)

[2] 福田 泰平，岩田 直樹，明田 修平，瀧本 栄二，川端 秀明， 半井 明大，窪田 歩，毛利 公一，Android における JDWP を利用した API 呼出し元モジュール特定手法

[3] Google Play，https://play.google.com/store