# Un análisis del software de Android preinstalado

Julien Gamba*†, Mohammed Rashed†, Abbas Razaghpanah‡, Juan Tapiador † y Narseo Vallina-Rodriguez*§

* IMDEA Networks Institute, † Universidad Carlos III de Madrid, ‡ Stony Brook University, § ICSI

*Resumen*

Android es un sistema operativo de código abierto. Esta característica hace posible que los fabricantes de dispositivos inteligentes desarrollen y desplieguen versiones personalizadas del sistema operativo junto con un conjunto de aplicaciones preinstaladas, a menudo para diferenciar sus productos en el mercado. Algunos proveedores de dispositivos han sido objeto de escrutinio por prácticas de recopilación de datos privados potencialmente invasivas y otros comportamientos dañinos o no deseados de las aplicaciones que pre-instalan en sus dispositivos Sin embargo, el panorama del software preinstalado en Android ha permanecido en gran parte inexplorado, particularmente en cuanto a las implicaciones de seguridad y privacidad de tales personalizaciones. Este artículo presenta el primer estudio a gran escala de software preinstalado en dispositivos Android fabricados por más de 200 vendedores. Nuestro trabajo se basa en un gran conjunto de datos de firmware de Android adquirido en todo el mundo utilizando métodos de crowdsourcing. Esto nos permite responder a preguntas relacionadas con las partes interesadas e involucradas en la cadena de suministro, desde fabricantes de dispositivos y operadores de redes móviles hasta organizaciones de terceros que ofrecen publicidad en línea y servicios de tracking (o rastreo) de usuarios. Nuestro estudio nos permite descubrir también las relaciones entre estos actores, y que parecen girar principalmente en torno a la publicidad y los servicios basados en datos. En general, nuestro estudio revela como la cadena de suministro en torno al código abierto de Android carece de transparencia y ha facilitado comportamientos potencialmente dañinos, incluyendo el acceso a datos sensibles sin consentimiento del usuario. Concluimos el artículo con una serie de recomendaciones para mejorar la transparencia, la atribución y responsabilidad en el ecosistema de Android

## I. Introducción

La naturaleza de código abierto de Android y su licencia de uso hace posible que cualquier fabricante desarrolle una versión personalizada del sistema operativo junto que, en muchos casos, incluye aplicaciones preinstaladas en la partición del sistema. La mayoría de los vendedores de teléfonos aprovechan esta oportunidad para dar valor añadido a sus productos como un diferenciador del mercado, generalmente a través de asociaciones y acuerdos con operadores de redes móviles (MNO), redes sociales y proveedores de contenido. Google no prohíbe este comportamiento, y ha desarrollado su compatibilidad con un programa de compatibilidad [7] que establece los requisitos que debe cumplir el sistema operativo modificado para seguir siendo compatible con las aplicaciones estándar de Android, independientemente de las modificaciones introducidas. Los dispositivos realizados por proveedores que forman parte del programa Android Certified Partners [4] vienen precargados con el conjunto de aplicaciones de Google (por ejemplo, Play Store y Youtube). Google no proporciona detalles sobre el proceso de certificación. Además, las empresas que desean incluir el servicio Google Play sin la certificación pueden externalizar el diseño del producto a un fabricante certificado (ODM) [6].

Certificado o no, no todo el software preinstalado no es deseado por los usuarios, y el término “bloatware” es utilizado a menudo. El proceso de cómo un conjunto particular de aplicaciones termina empaquetado en el firmware de un dispositivo no es transparente para el usuario final, y varios casos aislados sugieren que carece de mecanismos de control que garanticen que el firmware está libre de vulnerabilidades [23], [24] o aplicaciones potencialmente maliciosas y no deseadas. Por ejemplo, en BlackHat USA 2017, Johnson et al. [81], [46] describieron una potente puerta trasera presente en el firmware de varios modelos de teléfonos inteligentes Android, incluido el popular BLU R1 HD. En respuesta a esta divulgación, Amazon

eliminó los productos Blu de su catálogo [2]. La empresa Shanghai Adups Technology Co. Ltd. fue identificada como responsable por este incidente. El mismo informe también discutió la presencia de potenciales vulnerabilidades en servicios críticos del sistema. El troyano Triada también se ha encontrado recientemente integrado en el firmware de varios teléfonos inteligentes Android de baja gama [76], [65]. Otros casos de malware encontrados preinstalados han sido Loki (spyware y adware) y Slocker (ransomware), que fueron detectados en el firmware de varios teléfonos de alta gama [5].

Los teléfonos Android también juegan un papel clave en las prácticas de recopilación de datos a gran escala seguidas por muchos agentes de la economía digital, incluidas las empresas de publicidad y seguimiento o rastreo de usuarios. OnePlus ha estado bajo sospecha por recopilar datos sensibles y muy detallados del usuario (PII) en sus teléfonos inteligentes [54], [53], y también por ofrecer la capacidad de permitir el rooteado remoto del teléfono [52], [51]. En julio de 2018, el New York Times reveló la existencia de acuerdos secretos entre Facebook y fabricantes de dispositivos como Samsung [31] para recopilar datos privados de usuarios sin su conocimiento. Este incidente está siendo investigado por las autoridades federales de los Estados Unidos [32]. Además, los usuarios de países en desarrollo que carecen de leyes de protección de datos estrictas pueden correr un riesgo aún mayor. El Wall Street Journal ha revelado la presencia de un aplicación preinstalada que envía la ubicación geográfica de los usuarios y los identificadores del dispositivo a GMobi, una agencia de publicidad móvil que se ha asociado a actividades de fraude publicitario [13], [66]. Recientemente, la Comisión Europea expresó públicamente su preocupación por los fabricantes chinos como Huawei, alegando que estaban obligados a cooperar con los servicios nacionales de inteligencia mediante la instalación de puertas traseras en sus dispositivos [29].

*Objetivos de investigación y hallazgos*

Hasta donde sabemos, ningún estudio ha analizado sistemáticamente el vasto ecosistema de software preinstalado en dispositivos Android y sus implicaciones de privacidad y seguridad. Este ecosistema ha permanecido en gran parte inexplorado debido a la dificultad inherente para acceder a dicho software a escala y a través de varios fabricantes. Esta situación hace que tal estudio sea más relevante aún, ya que $i)$ estas aplicaciones, que normalmente no están disponibles en las tiendas de aplicaciones, han escapado al escrutinio de los investigadores y reguladores; y $ii)$ los usuarios finales desconocen su presencia en el dispositivo, lo que podría implicar falta de consentimiento en la obtención de los datos obtenidos. En este artículo, buscamos arrojar luz sobre la presencia y el comportamiento del software preinstalado en los dispositivos Android. En particular, nuestro objetivo es responder las siguientes preguntas:

- ¿Cuál es el ecosistema de aplicaciones preinstaladas en dispositivos Android así como los agentes que forman la cadena de suministro?
- ¿Cuáles son las relaciones entre los proveedores de dispositivos y otras partes interesadas (p. Ej., MNO y servicios de terceros de publicidad y rastreo de usuarios)?
- ¿Las aplicaciones preinstaladas ¿recopilan información privada y de identificación personal (PII)? Si es así, ¿con quién lo comparten?
- ¿Hay alguna aplicación dañina u otra potencialmente peligrosa entre el software preinstalado?

Para abordar estas preguntas, desarrollamos una agenda de investigación que gira en torno a cuatro elementos principales:

1. Recopilamos muestras de firmware así como información de tráfico de dispositivos del mundo real utilizando métodos de crowdsourcing (§II). Gracias a ello, hemos obtenido muestras de firmware de 2.748 usuarios que abarcan 1.742 modelos de dispositivos de 214 fabricantes. Nuestra base de usuarios cubre 130 países que incluyen los principales mercados de Android. Nuestro conjunto de datos contiene 424,584 muestras de software diferente, de los cuales sólo el 9 % de los APK están presentes en Google Play. Complementamos este conjunto de datos con flujos de tráfico asociados con 139,665 aplicaciones únicas, incluidas las preinstaladas, que han sido proporcionadas por más de 20.4K usuarios de la aplicación Lumen [85] de 144 países. Hasta donde sabemos, este es el mayor conjunto de datos con muestras de firmware de Android del mundo real analizado hasta ahora.
2. Realizamos una investigación del ecosistema de aplicaciones de Android preinstaladas y los actores involucrados en la cadena de suministro (§III) gracias al análisis del manifiesto de los ejecutables Android, sus firmas y certificados de desarrollador, y las librerías de terceros (TPL) que utilizan. Nuestro análisis cubre 1.200 desarrolladores únicos asociados con los principales

fabricantes, proveedores, operadores de redes móviles y empresas de servicios de Internet. Este análisis descubre un vasto panorama de librerías de terceros (11,665 TPL únicos), muchas de las cuales proporcionan principalmente servicios basados en la obtención y procesado de datos personales como la publicidad en línea, servicios de tracking y analytics, y redes sociales
3. Extrajimos y analizamos un amplio conjunto de permisos personalizados (4,845) declarados por proveedores de hardware, MNO y servicios de terceros que abarcan empresas de seguridad, alianzas y consorcios industriales, fabricantes de conjuntos de chips y desarrolladores de navegadores de Internet. Tales permisos pueden potencialmente exponer datos y servicios sensibles a otras aplicaciones para acceder a recursos privilegiados del sistema y datos confidenciales en una forma que elude el modelo de permisos de Android. Una inspección manual revela una cadena de suministro muy compleja que involucra diferentes agentes, y posibles asociaciones comerciales entre ellos (§IV).
4. Llevamos a cabo un análisis del comportamiento de casi el 50 % de las aplicaciones en nuestro conjunto de datos utilizando herramientas de análisis estático y dinámico.(§V). Nuestros resultados revelan que una parte importante del software preinstalado exhibe un comportamiento potencialmente dañino o no deseado. Si bien se sabe que la recopilación de datos personales y el seguimiento de usuarios es generalizada en el ecosistema de aplicaciones de Android en su conjunto [77], [83], [84] encontramos que también es bastante frecuente en aplicaciones preinstaladas. Hemos identificado instancias de actividades de seguimiento de usuarios mediante software de Android preinstalado, y bibliotecas de terceros integradas, que van desde la recopilación del conjunto habitual de PII y datos de geolocalización a prácticas más invasivas que incluyen el acceso al correo electrónico personal y metadatos sensibles como las llamadas telefónicas, contactos y una variedad de estadísticas de comportamiento y de uso en algunos casos. También encontramos algunas muestras de malware aisladas pertenecientes a familias de malware conocidas, según VirusTotal, con prevalencia en los últimos años (por ejemplo, Xynyin, SnowFox, Rootnik, Triada y Ztorg),y troyanos genéricos que muestran un conjunto estándar de comportamientos maliciosos (por ejemplo, promoción silenciosa de aplicaciones, fraude de SMS, fraude publicitario y fraude de clic de URL).

Con todo esto, nuestro trabajo revela por primera vez relaciones complejas entre los actores del ecosistema de Android, en los que los datos del usuario parecen ser el motor principal de esta industria. Descubrimos un vasto ecosistema de agentes involucrados en el desarrollo de software móvil, así como practicas de ingeniería de software bastante deficientes, así como una falta de transparencia en la cadena de suministro que aumenta innecesariamente los riesgos de seguridad y privacidad de los usuarios finales. Concluimos este documento con varias recomendaciones para paliar esta situación, incluidos modelos de transparencia para mejorar la atribución y responsabilidad de los agentes, y mecanismos más claros para obtener el consentimiento informado del usuario. Dada la escala del ecosistema y la necesidad de realizar inspecciones manuales para garantizar la seguridad de las muestras obtenidas, pondremos nuestro dataset a disposición de la comunidad investigadora y los reguladores gradualmente para impulsar otras investigaciones.

## II. Obtención de datos

La obtención de muestras de aplicaciones preinstaladas y otros artefactos de software (por ejemplo, certificados instalados en la root store del sistema) a escala es un reto desafiante. Esto se debe a que comprar todos los modelos de teléfonos móviles (y sus muchas variaciones) disponibles en el mercado es inviable. Por tanto, hemos decidimos hacer una herramienta que nos permita a través de crowdsourcing la colección de software preinstalado: Firmware Scanner [33]. Utilizando Firmware Scanner, obtuvimos software preinstalado de 1,742 modelos de dispositivos. También decidimos usar Lumen, una aplicación que tiene como objetivo promover la transparencia móvil y permitir el control del usuario sobre su tráfico móvil [85], [48]. Lumen nos permite correlacionar la información que extraemos del análisis estático, para un subconjunto de aplicaciones móviles, con tráfico de red generado por usuarios móviles mientras interactúan organicamente con sus dispositivos. En esta sección, explicamos los métodos de investigación y análisis implementados por cada herramienta y presentamos nuestro dataset. En el resto de esta sección, explicamos los métodos implementados por cada aplicación y

presentamos nuestros datasets. Discutimos las implicaciones éticas de nuestra recopilación de datos en la Sección II-C.

### *II-A. Firmware Scanner*

Disponible públicamente en Google Play [33], FirmwareScanner es una aplicación de Android diseñada para el propósito de este estudio. FirmwareScanner busca y extrae aplicaciones preinstaladas y archivos DEX en las carpetas `/system/`, bibliotecas `lib` y `lib64` ubicadas en `/system/`, cualquier archivo en la carpeta `/system/vendor/` (si ese directorio existe), y certificados raíz ubicados en `/system/etc/security/cacerts/`. Podemos distinguir las aplicaciones preinstaladas de las instaladas por el usuario ya que estas últimas son instaladas en `/data/app/`. Para reducir el tiempo de escaneo y subida, FirmwareScanner calcula los hash MD5 de los archivos relevantes (por ejemplo, aplicaciones, bibliotecas y certificados raíz) y luego envía la lista de estos hash a nuestro servidor. Solo aquellos ficheros que falten en nuestro dataset se suben a nuestro servidor a través de una conexión Wi-Fi con el fin de no afectar el plan de datos del usuario.

**Dataset:** Gracias a 2.748 usuarios que han instalado orgánicamente el Firmware Scanner, obtuvimos versiones de firmware para 1.742 modelos de dispositivos únicos[1] de 214 fabricantes[2] como se resume en la Tabla I.

Nuestro conjunto de datos contiene 424,584 archivos únicos (de acuerdo a su hash MD5) como se muestra en la Figura 1 para un subconjunto de fabricantes seleccionados. Para cada dispositivo trazamos tres puntos, uno para cada tipo de archivo, mientras que la forma indica la versión principal de Android instalada en el dispositivo.[3] Como podemos ver, el número de archivos preinstalados varía mucho de un vendedor a otro. Aunque no es sorprendente ver una gran cantidad de bibliotecas nativas debido a las diferencias de hardware, algunos de los proveedores incorporan cientos de aplicaciones adicionales (es decir, archivos ".apk") en comparación con otros fabricantes que ejecutan la misma versión de Android. El resto de nuestro estudio se centra en 82,501 aplicaciones de Android (es decir, ".apk") presentes en nuestro dataset, dejando el análisis de certificados raíz y bibliotecas para trabajos futuros.

Nuestra base de usuarios se distribuye geográficamente en 130 países, sin embargo, el 35 % de nuestros usuarios se encuentran en Europa, el 29 % en América (Norte y Sur), y 24 % en Asia. Además, hasta el 25 % y el 20 % del número total de dispositivos en nuestro conjunto de datos pertenecen a Samsung y Huawei, respectivamente. Estas figuras son coherentes con las estadísticas de mercado disponibles en la web [34], [9]. Mientras que ambos fabricantes son proveedores certificados por Google, nuestro conjunto de datos también contiene dispositivos Android de gama baja de fabricantes que se dirigen a mercados como Tailandia, Indonesia e India: muchos de estos proveedores no están certificados por Google. Finalmente, para evitar introducir sesgos en nuestro resultados, excluimos 321 teléfonos potencialmente rooteados.

### *II-B. Lumen*

Lumen es una aplicación de Android disponible en Google Play que tiene como objetivo promover la transparencia móvil y permitir el control del usuario sobre sus datos personales y tráfico. Lumen utiliza el permiso VPN de Android para interceptar y analizar todo el tráfico generado por el dispositivo en

[1] Utilizamos el hash MD5 del IMEI para identificar de forma exclusiva a un usuario, y la firma digital de compilación del proveedor para identificar de forma única un modelo de dispositivo determinado. Tenga en cuenta que dos dispositivos con la misma firma digital pueden personalizarse debido a la presencia de otros actores en la cadena de ensamblaje como operadores móviles y, por lo tanto, pueden tener diferentes aplicaciones preinstaladas.

[2] Confiamos en la información proporcionada por el fabricante que podría ser falsa. Por ejemplo, Alps renombra como ïPhone.ªlgunos de sus modelos Android. Según la información disponible en varios mercados online, son réplicas de iOS basadas en Android.

[3] Encontramos que 5,244 de las aplicaciones no implementan ninguna actividad, servicio o receptor Android (Activity, Service or Provider). Estas aplicaciones pueden usarse potencialmente como proveedores de recursos y contenedores (por ejemplo, imágenes, o fuentes) para otras aplicaciones.

[4] Consideramos que un dispositivo dado está rooteado de acuerdo a tres señales. Primero, cuando Firmware Scanner ha finalizado la carga de binarios preinstalados, la aplicación pregunta al usuario si el teléfono está rooteado de acuerdo con su propio entendimiento (tenga en cuenta que el usuario puede optar por no responder la pregunta). Como complemento, utilizamos la biblioteca RootBeer [62] para verificar de forma programática si un dispositivo está potencialmente rooteado. Si alguna de estas fuentes indica que el dispositivo está potencialmente rooteado, lo consideramos como tal. Finalmente, descartamos dispositivos donde hay evidencia de que se han instalado ROM personalizadas (por ejemplo, LineageOS). Discutimos las limitaciones de este método en la Sección VI.

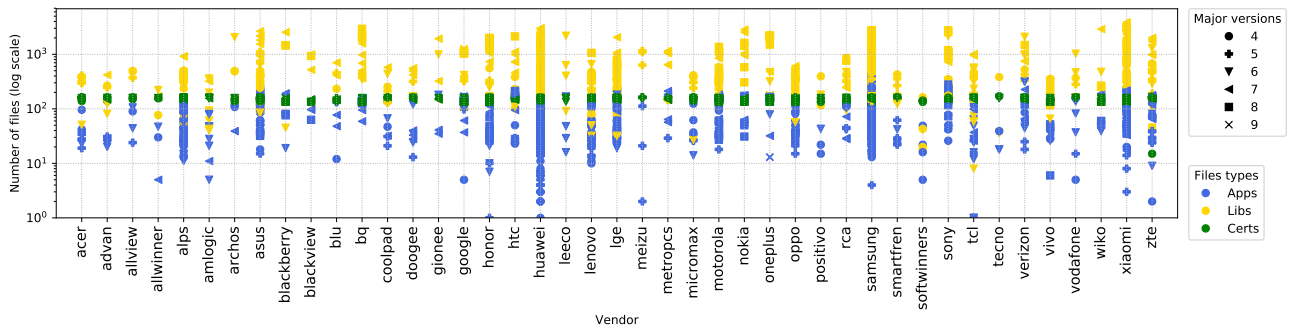


Figura 1: Número de archivos por proveedor. No mostramos los proveedores para los que tenemos menos de 3 dispositivos

| Vendedor | País | Proveedor certificado | Firmas digital de compilación | Usarios | Archivos (med.) | Apps (med.) | Bibs. (med.) | DEX (med.) | Certificados (med.) | Archivos (total) | Apps (total) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Samsung** | South Korea | Yes | 441 | 924 | 868 | 136 | 556 | 83 | 150 | 260,187 | 29,466 |
| **Huawei** | China | Yes | 343 | 716 | 1,084 | 68 | 766 | 96 | 146 | 150,405 | 12,401 |
| **LGE** | South Korea | Yes | 74 | 154 | 675 | 84 | 385 | 89 | 150 | 58,273 | 3,596 |
| **Alps Mobile** | China | No | 65 | 136 | 632 | 56 | 385 | 46 | 148 | 29,288 | 2,883 |
| **Motorola** | US/China | Yes | 50 | 110 | 801 | 127 | 454 | 62 | 151 | 28,291 | 2,158 |
| **Total (214 vendedores)** | — | **22 %** | **1,742** | **2,748** | | | | | | **424,584** | **82,501** |

Cuadro I: Estadísticas generales para los 5 principales proveedores de nuestro conjunto de datos.

espacio de usuario e *in situ*, incluso si está cifrado o encriptado, sin necesidad de rootear el dispositivo. Al ejecutarse localmente en el dispositivo del usuario, Lumen es capaz de correlacionar los flujos de tráfico con la información a nivel del sistema y la actividad de la aplicación. La arquitectura de Lumen está disponible públicamente [85]. Lumen nos permite determinar con precisión qué aplicación es responsable de una fuga de datos personales observada desde el punto de vista del usuario y desencadenado por usuarios reales en condiciones normales de uso. Como todo el análisis ocurre en el dispositivo, solo los metadatos de tráfico procesados se extraen del dispositivo.

**Dataset:** Para este estudio, utilizamos registros de tráfico anónimos proporcionados por más de 20.4K usuarios de Lumen de 144 países (según estadísticas de la Play Store) procedentes de teléfonos Android fabricados por 291 proveedores. Esto incluye 34M de flujos de tráfico de más de 139K aplicaciones únicas (298K combinaciones únicas de nombre y versión de paquete o aplicación). Sin embargo, como Lumen no recopila la firma digital de la aplicación, para encontrar la superposición entre el conjunto de datos de Lumen y las aplicaciones preinstaladas, relacionamos los datasets (es decir, las based de datos de Lumen y FirmwareScanner) a través del paquete Android, versión de aplicación y proveedor de dispositivo o fabricante. Mientras este método no garantiza que las aplicaciones superpuestas sean exactamente las mismas, es seguro asumir que los teléfonos que no están rooteados no tendrán diferentes aplicaciones con los mismos nombres de paquete y versiones de la aplicación. Como resultado, tenemos 1.055 combinaciones únicas de aplicación/versión/vendedor de dispositivo presentes en ambos datasets.

### *II-C. Preocupaciones éticas*

Nuestro estudio implica la recopilación de datos de usuarios reales que instalaron orgánicamente FirmwareScanner o Lumen en sus dispositivos. Por lo tanto, seguimos los principios del consentimiento informado [75] y evitamos la recopilación de cualquier información personal o sensible. Buscamos la aprobación de nuestra Comité de Ética institucional y de nuestro Oficial de Protección de Datos (DPO) antes de comenzar con la recolección y procesado de los datos. Ambas herramientas también proporcionan amplias y detalladas políticas de privacidad en su perfil de Google Play. A continuación discutimos detalles y consideraciones específicos para cada herramienta.

**Firmware Scanner:** Esta aplicación recopila algunos metadatos sobre el dispositivo para atribuir observaciones a los fabricantes (por ejemplo, su modelo y huella digital o firma) junto con algunos datos sobre las aplicaciones preinstaladas (extraídas del

Administrador de paquetes), operador de telefonía móvil (MNO) y usuario (la zona horaria y los códigos MCC y MNC de su tarjeta SIM, si están disponibles para identificar aspectos geográficos). Calculamos el hash MD5 del IMEI del dispositivo para identificar duplicados y versiones de firmware actualizadas para un dispositivo determinado.

**Lumen:** los usuarios de Lumen deben optar activamente para participar en el experimento dos veces antes de iniciar la intercepción de tráfico [75]. Lumen preserva la privacidad de sus usuarios al realizar el procesamiento y análisis de flujo en el dispositivo, solo enviando metadatos de tráfico generado por las apps anóninimamente para fines de investigación. Lumen no recopila ningún identificador de dispositivo o de usuario, ni el contenido de las conexiones generadas por el dispositivo del usuario. Para proteger aún más la privacidad del usuario, Lumen también ignora todos los flujos generados por aplicaciones sensibles como el navegador web, ya que estas pueden potencialmente desanonimizar a un usuario. Lumen permite al usuario deshabilitar la intercepción de tráfico en cualquier momento.

## III. Análisis de la Cadena de Ensamblado

La apertura del sistema operativo Android ha permitido la creación de una compleja cadena de suministro formada por diferentes tipos de agentes y organizaciones, ya sean fabricantes, operadores móviles, desarrolladores y organizaciones afiliadas, y distribuidores. Estos actores pueden agregar aplicaciones y características propietarias a dispositivos Android con el fin de brindar una mejor experiencia de usuario, dar valor añadido a sus productos o brindar acceso a servicios propietarios. Sin embargo, tras estas prácticas también podría haber un beneficio económico (mutuo) basado en la obtención y procesado de datos personales [31], [13]. Esta sección proporciona una descripción general del software Android preinstalado para descubrir algunas de las áreas grises que rodean a estas prácticas, analizar el amplio y diverso conjunto de desarrolladores involucrados, la presencia de bibliotecas de tracking y publicidad de terceros, y analizar el papel de cada agente.

### *III-A. Ecosistema de desarrolladores*

Comenzamos nuestro estudio analizando las organizaciones que firman cada aplicación preinstalada. Primero, agrupamos las aplicaciones por los certificados utilizados para firmar el software, y luego confiamos en la información presente en el campo `Issuer` del certificado para identificar el desarrollador [14]. A pesar de que esta es la señal más fiable para identificar a la organización que firma el software, sigue siendo ruidosa ya que una empresa puede usar múltiples certificados, uno para cada unidad organizativa. Más importante aún, estos son certificados autofirmados, lo que reduce significativamente la confianza que se puede depositar en ellos.

Como resultado de estas limitaciones, no pudimos identificar a la compañía detrás de varios certificados (indicados como “Compañía desconocida” en la Tabla II) debido a información insuficiente o dudosa en el certificado: por ejemplo, el campo `Issuer` de muchas apps preinstaladas sólo contiene las menciones de la Compañía y Departamento. Además, hemos encontrado aplicaciones firmadas por 42 certificados diferentes de *.Android Debug”* en teléfonos de 21 marcas diferentes. Estos casos reflejan prácticas de desarrollo pobres y potencialmente inseguras a medida que se utiliza el certificado de depuración de Android para firmar automáticamente aplicaciones en entornos de desarrollo, permitiendo así que otras aplicaciones firmadas con ese certificado accedan a su funcionalidad sin solicitar ningún permiso. La mayoría de las tiendas de aplicaciones (incluida Google Play) no aceptarán la publicación de un aplicación firmada con un certificado de depuración [8]. Además, también encontramos hasta 115 certificados que sólo mencionan *“Android”* en el campo del emisor. Se supone que una gran parte (43 %) de esos certificados se emiten en los EE. UU., Mientras que otros parecen haber sido emitidos en Taiwán (16 %), China (13 %) y Suiza (13 %). En ausencia de una lista pública de certificados oficiales de desarrollador, o de una autoridad certificadora, no es posible verificar su autenticidad o conocer a su propietario, como se discutirá en la Sección VI.

Con estas limitaciones en mente, extrajimos 1,200 certificados y firmas únicos de nuestro conjunto de apps preinstaladas. La tabla II muestra las 5 empresas más presentes en el caso de los vendedores de teléfonos (izquierda) y otro tipo de empresas (derecha). Este análisis descubrió un vasto ecosistema de software de terceros, incluidas grandes empresas digitales (por ejemplo, LinkedIn, Spotify y TripAdvisor), así como servicios de publicidad y tracking. Este es el caso de ironSource, una empresa de publicidad que firma software preinstalado [42] que ha sido encontrado en

| Nombre de empresa | Cantidad de certificados | País | Proveedor certificado? |
|---|---|---|---|
| Google | 92 | United States | N/A |
| Motorola | 65 | US/China | Yes |
| Asus | 60 | Taiwan | Yes |
| Samsung | 38 | South Korea | Yes |
| Huawei | 29 | China | Yes |
| **Total (vendedores)** | **740** | — | — |

| Nombre de empresa | Cantidad de certificados | País | Cantidad de vendedores |
|---|---|---|---|
| MediaTek | 19 | China | 17 |
| Aeon | 12 | China | 3 |
| Tinno Mobile | 11 | China | 6 |
| Verizon Wireless | 10 | United States | 5 |
| *Unknown company* | 7 | China | 1 |
| **Total** | **460** | — | **214** |

Cuadro II: **Izquierda:** los 5 desarrolladores más frecuentes (según el número total de aplicaciones firmadas por ellos), y **derecha:** para otros empresas.

Asus, Wiko y otros fabricantes de dispositivos, y TrueCaller, un servicio para bloquear llamadas o mensajes de texto no deseados [56]. Según su sitio web y también fuentes independientes [39], [70], TrueCaller utiliza mecanismos de crowdsourcing para construir un gran conjunto de datos de números de teléfono utilizado para spam y también para publicidad. Del mismo modo, hemos encontrado 123 aplicaciones (por su MD5) firmadas por Facebook. Estas aplicaciones se encuentran en 939 dispositivos, el 68 % de los cuales son de Samsung. También hemos encontrado aplicaciones firmadas por AccuWeather, un servicio meteorológico que ha sido previamente denunciado por obtener datos personales agresivamente [86]. Finalmente, hemos encontrado otras compañías de software como Adups, responsable de la puerta trasera de Adups [45], y GMobi [35], una compañía de publicidad móvil previamente acusada de prácticas dudosas por el Wall Street Journal [13].

### *III-B. Servicios de terceros*

Al igual que en la web y apps disponibles en Google Play Store, los desarrolladores de aplicaciones móviles pueden incrustar en su software preinstalado bibliotecas de terceros (TPL, o third-party libraries) proporcionadas por otros empresas, incluidas las bibliotecas (SDK) proporcionadas por servicios de publicidad online, servicios de analíticas y tracking, o redes sociales. En esta sección usamos LibRadar, una herramienta resistente a la ofuscación, para identificar los TPL utilizados en las aplicaciones de Android [90], en nuestro conjunto de datos para examinar su presencia y las posibles implicaciones sobre la privacidad para los usuarios: cuando están presentes en aplicaciones preinstaladas, las TPL tienen la capacidad de monitorear actividades del usuario longitudinalmente [89], [84]. Excluimos las TPL conocidas que brindan soporte de desarrollo como biblioteca de protocolos de red. Primero, clasificamos las 11,665 TPL únicas identificadas por LibRadar++ de acuerdo con las categorías definidas por [82], AppBrain [50], y PrivacyGrade [57]. Clasificamos manualmente aquellas TPLs que no fueron clasificadas previamente.

| Categoría | # bibliotecas | # apps | # vend. | Ejemplo |
|---|---|---|---|---|
| Anuncio | 164 (107) | 11,935 | 164 | Braze |
| Analitica móvil | 100 (54) | 6,935 | 158 | Apptentive |
| Redes sociales | 70 (20) | 6,652 | 157 | Twitter |
| **Todas las categorias** | **334** | **25,333** | **165** | — |

Cuadro III: Categorías de TPL seleccionadas presentes en aplicaciones preinstaladas. Entre paréntesis, reportamos el número de TPLs identificadas cuando se agrupan por el nombre del Java package.

Una vez clasificadas, enfocamos nuestro estudio en aquellas categorías que podrían causar daño a la privacidad de los usuarios, como las librearias asociadas a servicios de tracking y la publicidad en línea. Encontramos 334 TPLs en tales categorías, como se resume en la Tabla III. Podríamos identificar empresas de la publicidad y seguimiento de usuarios como Smaato (especializada en anuncios geográficos [63]), GMobi, Appnext, ironSource, Crashlytics y Flurry. Algunos de estos proveedores externos también se han identificado en la Sección III-A ya que también despliegan sus propias apps como software preinstalado gracias a acuerdos con los vendedores, o son desarrolladores destacados responsables de aplicaciones publicadas en Google Play Store [84]. Encontramos 806 aplicaciones que incorporan el Facebook Graph SDK que se distribuye a través de 748 dispositivos. Los certificados de estas aplicaciones sugieren que 293 de ellas fueron firmadas por el proveedor del dispositivo y 30 por operadores de telefonía móvil (sólo 98 están firmadas por el propio Facebook). La presencia de los SDK de Facebook en aplicaciones preinstaladas podría, en algunos casos, ser explicado por los

acuerdos comerciales establecidos por Facebook con proveedores de Android como reveló el New York Times [31].

Encontramos otras compañías digitales que ofrecen servicios de analítica y esquemas de monetización de aplicaciones basados en datos como Umeng, Fyber (anteriormente Heyzap) y Kochava [84]. Además, también encontramos casos de empresas de análisis avanzado en teléfonos Asus como Appsee [16] y Estimote [27]. Según sus sitio web, Appsee es un TPL que permite a los desarrolladores registrar eventos relacionados con la interacción del usuario con su dispositivo [15], incluidos los eventos táctiles y grabar el contenido de la pantalla [83]. Si, por sí solo, grabar la pantalla del usuario no constituye una fuga de privacidad,registrar y cargar estos datos podría filtrar involuntariamente información privada, como detalles de autenticación y credenciales del usuario. Estimote proporciona soluciones avanzadas para geo-localización en interiores [27]. El SDK de Estimote permite que una aplicación reaccione a balizas inalámbricas virtuales con el fin, por ejemplo, de enviar notificaciones personalizadas y publicidad al usuario al ingresar físicamente en una tienda.

Finalmente, encontramos TPL proporcionados por compañías especializadas en el mercado chino [90] en 548 aplicaciones preinstaladas. Lo mas relevantes son el SDK de Tencent, AliPay (un servicio de pago) y el SDK de Baidu [19] (para publicidad y geolocalización / geocodificación servicios). Los dos últimos son posiblemente utilizados como alternativa a Google Pay y Maps en el mercado chino, respectivamente. Solo una de las aplicaciones que incorporan estos SDK está firmada por el proveedor de servicios real, lo que indica que su presencia como aplicaciones preinstaladas probablemente se deba a las decisiones de diseño de los los propios fabricantes y sus desarrolladores o compañías afiliadas.

### *III-C. Aplicaciones disponibles públicamente*

Analizamos la Google Play Store para identificar cuántas de las aplicaciones preinstaladas encontradas por Firmware Scanner están disponibles para el publico en los mercados de aplicaciones. Este análisis se llevó a cabo el 19 de noviembre de 2018 y sólo usamos el nombre del paquete de la app preinstalada como parámetro. Descubrimos que solo el 9 % de los paquetes en nuestro dataset están indexados en Google Play Store. Para aquellos indexados, pocas categorías dominan el espectro de aplicaciones preinstaladas de acuerdo con los metadatos de Google Play, en particular comunicación, entretenimiento, productividad, herramientas y aplicaciones multimedia

La baja presencia de aplicaciones preinstaladas en Google Play Store sugiere que este tipo de software podría haber escapado a cualquier escrutinio previo. De hecho, hemos encontrado ejemplos de aplicaciones preinstaladas desarrolladas por organizaciones prominentes que no son publicadas en Google Play. Por ejemplo, software desarrollado y firmado por Facebook (véase, `com.facebook.appmanager`), Amazon y CleanMaster entre otros. Del mismo modo, encontramos versiones no públicas de navegadores web populares (por ejemplo, el Navegador UME, y Opera).

De acuerdo a la información disponible en la Play Store sobre la versión de los paquetes preinstalados, encontramos que el software preinstalado disponible públicamente en la Play Store se actualizan con más frecuencia que el resto de las aplicaciones preinstaladas: el 74 % de las aplicaciones no públicas no parece actualizarse y el 41 % de ellos permanecieron sin parchear durante 5 años o más. Si existe una vulnerabilidad en uno de estos aplicaciones (consulte la Sección V), el usuario puede estar en riesgo mientras siga usando el dispositivo.

## IV. Análisis de Permisos

Android implementa un modelo de permisos para controlar el acceso de las aplicaciones a datos confidenciales, sensibles, y recursos del sistema [55]. Por defecto, no se permite que las aplicaciones realicen ninguna operación protegida. Los permisos de Android no se limitan a los definidos por AOSP: cualquier desarrollador de aplicaciones, incluidos los fabricantes, puede definir sus propios permisos personalizados (“custom permissions”) para exponer su funcionalidad a otros aplicaciones [25]. Utilizamos Androguard [3] para extraer y estudiar los permisos, tanto declarados como solicitados, por las aplicaciones preinstaladas. Nos centramos principalmente en los permisos personalizados, ya que $i)$ los servicios preinstalados tienen acceso privilegiado a los recursos del sistema, y $ii)$ los servicios preinstalados privilegiados pueden (involuntariamente) exponer servicios y datos críticos, incluso sin requerir el acceso a los permisos oficiales de Android.

### *IV-A. Permisos personalizados declarados*

Identificamos 1.795 nombres únicos de paquetes de Android en 108 proveedores de Android que definen 4.845 permisos personalizados. Excluimos los permisos definidos por AOSP y los asociados con la mensajería en la nube de Google (GCM) [36]. El número de permisos declarados por fabricante de dispositivos Android varía según las marcas y los modelos debido a las acciones de otros agentes presentes en la cadena de suministro. Clasificamos las organizaciones que declaran permisos personalizados en 8 grupos como se muestra en la Tabla IV: fabricantes de dispositivos inteligentes (p. Ej., Samsung), MNO (p. Ej., Verizon), servicios de terceros (p. Ej., Facebook), empresas de AV (p. Ej., Avast), alianzas industriales (p. Ej., GSMA), fabricantes de hardware y chipsets (p. Ej., Qualcomm) y navegadores (p. Ej., Mozilla). No pudimos identificar con certeza las organizaciones responsables para el 9 % de los permisos personalizados identificados. La ausencia de buenas prácticas de desarrollo entre los desarrolladores y fabricantes complica esta clasificación [25], obligándonos a seguir un proceso semi-manual que implica analizar múltiples señales para identificar sus posibles propósitos y la atribución.

Como se muestra en la Tabla IV, el 31 % de todos los permisos personalizados declarados están definidos por 31 proveedores de teléfonos de acuerdo con nuestra clasificación. La mayoría de ellos están asociados con servicios proprietarios como soluciones de administración de dispositivos móviles (MDM) para clientes corporativos. Sin embargo, tres proveedores representan más del 68 % de los permisos personalizados totales; Samsung (41 %), Huawei(20 %) y Sony (anteriormente Sony-Ericsson, 7 %). La mayoría de los permisos personalizados definidos por los proveedores de hardware, junto con el conjunto de permisos declarados por fabricantes de chips y MNOs están expuestos por los servicios principales y críticos del framework Android, incluido el navegador predeterminado `com.android.browser`. Desafortunadamente, como se demostró en el caso MediaTek [78], exponer tales recursos sensibles en servicios críticos puede aumentar la superficie de ataque si no se implementan con cuidado.

Un análisis exhaustivo de los permisos personalizados también sugiere (y en algunos casos confirma) la presencia de acuerdos comerciales entre proveedores de teléfonos móviles, operadores móviles, servicios de análisis de datos y analíticas (por ejemplo, Baidu, ironSource y Digital Turbine),y grandes corporaciones de la industria digital (por ejemplo, Skype, LinkedIn, Spotify, CleanMaster y Dropbox). También encontramos permisos personalizados asociados con módulos previamente identificados como vulnerables (por ejemplo, MediaTek) y servicios potencialmente dañinos (por ejemplo, Adups). Discutimos un número selecto de casos de interés a continuación.

**Soluciones VPN:** Android proporciona soporte nativo a clientes VPN de terceros. Esta característica se considera tan sensible ya que la aplicación que solicite este permiso adquiere la capacidad de romper el sandboxing de Android y monitorear el tráfico de los usuarios, independientemente de la aplicación [67], [79]. El análisis de los permisos personalizados revela que Samsung y Meizu implementan su propio servicio VPN. No está claro por qué estas implementaciones propietarias de la interfaz VPN existen, pero varios desarrolladores de soluciones VPN han reportado problemas de compatibilidad ya que el servicio VPN predeterminado de Android no está soportando en dichos teléfonos [1], [85], [79]. Un análisis completo de estos servicios propietarios de VPN se propone como trabajo futuro.

**Facebook:** Encontramos 6 paquetes diferentes firmados por Facebook, tres de ellos no disponibles en Google Play. Estos paquetes declaran 18 permisos personalizados como se muestra en la Tabla V y que se han encontrado en 24 proveedores de Android, incluidos Samsung, Asus, Xiaomi, HTC, Sony y LG. Según las quejas de los usuarios, dos de estos paquetes (`com.facebook.appmanager` y `com.facebook.system`) parecen descargar automáticamente y de forma silenciosa software de Facebook como Instagram en los teléfonos de los usuarios [68], [69]. También encontramos interacciones entre Facebook y MNOs como Sprint.

**Baidu:** El permiso para obtener la ubicación geográfica de Baidu está expuesto por aplicaciones preinstaladas, incluidos los módulos principales de Android, en 7 diferentes vendedores, principalmente chinos. Este permiso parece estar asociado con la API de geo-codificación y geo-localización de Baidu [18] y podría permitir a desarrolladores de apps el acceso a datos sensibles de usuario, eludiendo la solicitud del permiso de ubicación de Android.

**Digital Turbine:** Hemos identificado 8 permisos personalizados en 8 servicios asociados con Digital Turbine y su subsidiaria LogiaGroup. Su política de privacidad indica que esta empresa recopila datos

| | Permisos | Proveedores | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | personalizados | Vendedor | Terceros | MNO | Chipset | AV / Securidad | Ind. Alliance | Navegador | Otro |
| **Total** | 4,845 (108) | 3,760 (37) | 192 (34) | 195 (15) | 67 (63) | 46 (13) | 29 (44) | 7 (6) | 549 (75) |
| **Módulos Android** | | | | | | | | | |
| `android` | 494 (21) | 410 (9) | — | 12 (2) | 4 (13) | — | 6 (7) | — | 62 (17) |
| `com.android.systemui` | 90 (15) | 67 (11) | 1 (2) | — | — | — | — | — | 22 (8) |
| `com.android.settings` | 87 (16) | 63 (12) | — | 1 (1) | — | — | — | — | 23 (8) |
| `com.android.phone` | 84 (14) | 56 (9) | — | 5 (2) | 3 (5) | — | — | — | 20 (10) |
| `com.android.mms` | 59 (11) | 35 (10) | — | 1 (2) | — | — | 1 (1) | — | 22 (8) |
| `com.android.contacts` | 40 (7) | 32 (3) | — | — | — | — | — | — | 8 (5) |
| `com.android.email` | 33 (10) | 18 (4) | — | — | — | — | — | — | 15 (17) |

Cuadro IV: Resumen de permisos personalizados por categoría de proveedor y su presencia en módulos básicos sensibles de Android seleccionados.El valor entre paréntesis informa el número de proveedores de Android en los que se encontraron permisos personalizados

| Paquete | Públic | # Fabricantes | # Permisos |
|---|---|---|---|
| `com.facebook.system` | No | 18 | 2 |
| `com.facebook.appmanager` | No | 15 | 4 |
| `com.facebook.katana` (Facebook) | Si | 14 | 8 |
| `com.facebook.orca` (Messenger) | Si | 5 | 5 |
| `com.facebook.lite` (FB Lite) | Si | 1 | 1 |
| `com.facebook.pages.app` | No | 1 | 4 |
| **Total** | 3 | 24 | 18 |

Cuadro V: Paquetes de Facebook preinstalados en teléfonos Android.

personales que van desde UID hasta registros de tráfico que podrían ser compartido con sus socios comerciales [26]. Según la información de SIM de estos dispositivos, los módulos preinstalados de Digital Turbine se encuentran principalmente en usuarios norteamericanos y asiáticos. Un nombre de paquete, `com.dti.att` ("dti"significa Digital Turbine Ignite), sugiere acuerdos comerciales con el operador de telefonía AT&T. Al inspeccionar su código fuente, este paquete parece implementar un servicio integral de administración de software, incluyendo las instalaciones y la eliminación de aplicaciones por parte de los usuarios que se rastrean y vinculan con datos personales, que sólo parecen estar "enmascarados" (es decir, hash-ed).

**ironSource:** la empresa de publicidad ironSource expone permisos personalizados relacionados con su AURA Enterprise Solutions [43]. Hemos identificado varios paquetes específicos de proveedores que exponen permisos personalizados de ironSource en dispositivos fabricados por proveedores como Asus, Wiko y HTC (el nombre del paquete y las firmas del certificado sugieren que esos módulos posiblemente se introducen con colaboración del vendedor). Según material comercial publicado por ironSource [44], AURA tiene acceso a más de 800 millones de usuarios por mes, y proporcionan acceso a servicios de análisis avanzados así como la precarga de software en los dispositivos de los clientes. Un análisis superficial de algunos de estos paquetes (por ejemplo, `com.ironsource.appcloud.oobe.htc`, `com.ironsource.appcloud.oobe.asus`) revela que proporcionan servicios Out-of-the-Box (OOBE) que permiten una precarga de software personalizada a cada usuario cuando estos abren su dispositivo por primera vez [43], al mismo tiempo que monitoran las actividades.

**Otros servicios de publicidad y seguimiento:** Analizar individualmente cada permiso personalizado introducido por servicios de terceros requeriría un análisis más allá del alcance de este estudio. Sin embargo, hay varios casos anecdóticos de interés que merecen ser considerados. Uno es el caso de una aplicación preinstalada firmada por Vodafone (Grecia) y presente en un dispositivo Samsung que expone un permiso asociado con Exus [30], una empresa especializada en gestión de riesgo crediticio y soluciones bancarias. Otro caso son los permisos presentes en algunos teléfonos Samsung y LG (probablemente vendidos por Verizon) relacionados con la empresa Synchronoss. Su política de privacidad reconoce la recopilación, el procesamiento y la compartición de datos personales [64].

**Servicios de protección de llamadas:** Identificamos tres compañías externas que brindan servicios para bloquear llamadas y mensajes de texto provenientes de teléfonos no deseados: Hiya [37], TrueCaller [56], y PrivacyStar [58]. La solución de Hiya parece estar integrada por T-Mobile (EE. UU.), Orange (España) y AT&T (EE. UU.) en teléfonos Samsung y LG, posiblemente subsidiados por las compañías, de acuerdo con los certificados usados para firmar la app. Las políticas de privacidad de Hiya y TrueCaller indican que recopilan datos personales de los usuarios, incluidos los contactos

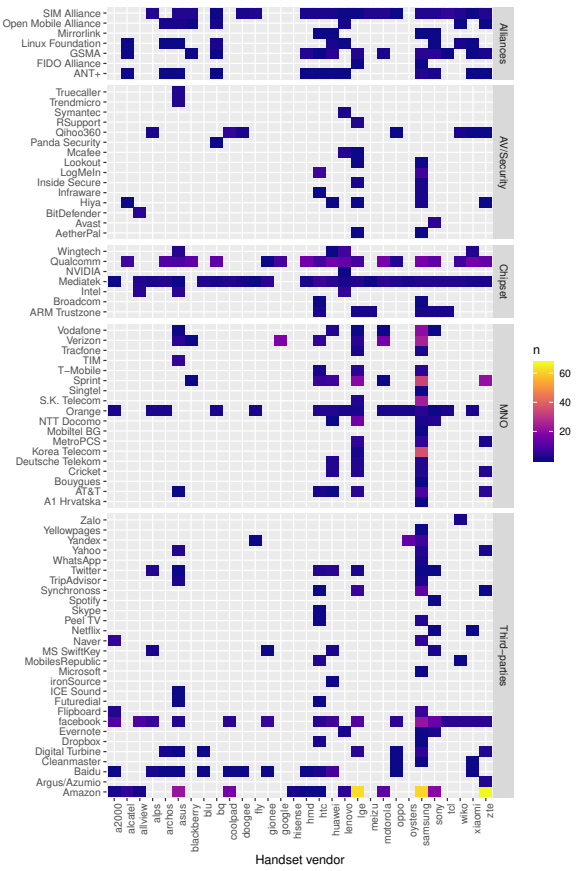


Figura 2: Permisos personalizados definidos por compañías de AV, MNOs, fabricantes de chipsets, y servicios de publicidad y tracking requeridos por apps preinstaladas.

almacenados en el dispositivo, varios UIDs y otra información personal [38][1]. La política de privacidad de PrivacyStar, en cambio, afirma que cualquier información recopilada de los contactos de un usuario dado “NO se exporta fuera de la aplicación para ningún propósito” [59].

### *IV-B. Permisos personalizados usados*

El uso de permisos por parte de aplicaciones de Android preinstaladas sigue una distribución potencial: 4.736 de las apps preinstaladas solicita al menos un permiso y 55 aplicaciones solicitan más de 100. El hecho de que las aplicaciones preinstaladas soliciten muchos permisos para implementar y ofrecer su servicio no implica necesariamente una violación de la privacidad del usuario. Sin embargo, identificamos un número significativo de apps que parecen potencialmente sobre-privilegiadas con actividades sospechosas como `com.jrdcom.Elabel` — una app firmada por TCLMobile que solicita 145 permisos y ha sido etiquetada como maliciosa por Hybrid Analysis —y `com.cube26.coolstore` (144 permisos). Del mismo modo, la calculadora preinstalada que se encuentra en un Xiaomi Mi 4c solicita la ubicación del usuario y el estado del teléfono, lo que le da acceso a identificadores unicos (UIDs) como el IMEI. Discutimos más instancias de aplicaciones con privilegios excesivos en la Sección V-C.

[1]Nota: la información presentada en su política de privacidad difiere cuando se accede desde una máquina en la U.E. o los EE. UU. En enero de 2019, ninguna de estas compañías mencionan la nueva directiva europea GDPR en sus políticas de privacidad.

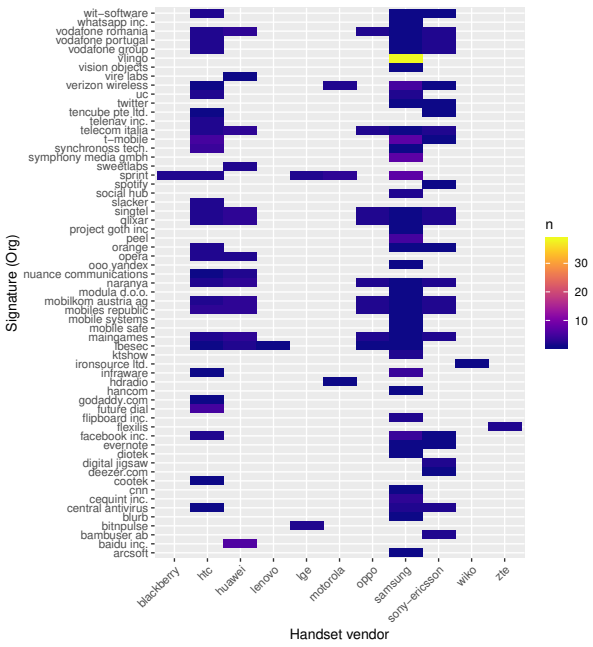


Figura 3: Apps accediendo a permisos propietarios definidos por fabricantes de dispositivos.

**Permisos peligrosos de Android.** Estadisticamente, una aplicación Android preinstalada solicita tres permisos peligrosos de AOSP (mediana). Cuando observamos el conjunto de permisos solicitados por una aplicación determinada (por el nombre del paquete) a través de todos los proveedores, podemos notar importantes diferencias. Investigamos tales variaciones en un subconjunto de 150 apps presentes al menos en 20 proveedores diferentes. Esta lista contiene principalmente servicios centrales de Android, así como aplicaciones firmadas por compañías independientes (p. Ej., Adups) y fabricantes de conjuntos de chips(p. Ej., Qualcomm).

Posteriormente, agrupamos todos los permisos solicitados por un nombre de paquete dado en

todos los modelos de dispositivos para cada marca. Como en el caso de los permisos personalizados expuestos, podemos ver una tendencia a privilegiar en exceso estos módulos en fabricantes específicos. Por ejemplo, el número de permisos solicitados por el módulo principal `android` puede variar de 9 permisos en un dispositivo Android ofrecido por Google — p. Ej. un Nexus 5 o un Pixel — a más de 100 en la mayoría de los dispositivos Samsung. Del mismo modo, mientras que el paquete `com.android.contacts` suele solicitar 35 permisos (mediana) en nuestra base de datos, este número supera los 100 para dispositivos Samsung, Huawei, Advan y LG.

**Permisos personalizados.** 2.910 aplicaciones preinstaladas solicitan al menos un permiso personalizado. El mapa de calor en la Figura 2 muestra la cantidad de permisos personalizados solicitados por paquetes preinstalados en un conjunto cuidadosamente seleccionado de fabricantes populares de Android (eje x). Como podemos ver, el uso de permisos personalizados también varía entre los proveedores. Aquellos asociados con terceros, sobre todo agentes destacados de la industria y la economía de los datos (p. Ej., Facebook), MNOs (p. Ej., Vodafone) y los servicios de AV / seguridad (p. Ej., Hiya) son los más los solicitados.

Este análisis revela posibles acuerdos comerciales más allá de los revelados en las secciones anteriores. Identificamos el desarrollador gracias a la firma de los apps que acceden a los permisos de ironSource, Hiya y AccuWeather. Este estado de cosas potencialmente permite que terceros puedan obtener acceso a permisos protegidos solicitados por otros paquetes preinstalados firmados con la misma firma. Además, encontramos paquetes firmados por Sprint que se parecen a los APKs de Facebook y Messenger de Facebook (`com.facebook.orca.vpl` y `com.facebook.katana.vpl`) y que además solicitan permisos relacionados con Flurry (un servicio de seguimiento en línea y analyticas propiedad de Verizon).

Las relaciones comerciales entre servicios de terceros y proveedores parecen ser bidireccionales como se muestra en la Figura 3. Esta figura muestra evidencia de 87 aplicaciones que acceden a los permisos de otros proveedores, incluidos los paquetes firmados por Facebook, ironSource, Hiya, Digital Turbine, Amazon, Verizon, Spotify, varios navegadores y MNOs. Estos están agrupados por la firma del desarrollador para fines de claridad. Como indica el mapa de calor, Samsung, HTC y Sony son los proveedores que habilitan la mayoría de los permisos personalizados solicitados por aplicaciones, posiblemente debido a su importante presencia en el mercado Android. Encontramos instancias de aplicaciones disponibles en Play Store que también solicitan dichos permisos. Desafortunadamente, estos permisos personalizados no se muestran a los usuarios cuando compran aplicaciones móviles en la tienda; por lo tanto, aparentemente se solicitan sin consentimiento, lo que puede causar graves daños a la privacidad de los usuarios cuando las aplicaciones las utilizan de forma incorrecta para obtener acceso a datos y recursos de sistema sensibles.

### *IV-C. Uso de permisos por TPLs*

Analizamos los permisos utilizados por las aplicaciones que incorporan al menos un TPL (librería de terceros). Estudiamos el acceso a permisos con una protección nivel de `signature` o `signature|privileged`, ya que solo pueden otorgarse a las aplicaciones del sistema [49] o a las firmadas con una firma del sistema. La presencia de TPLs en aplicaciones preinstaladas que solicitan acceso a una firma o permiso peligroso por lo tanto, puede darle acceso a recursos muy sensibles sin el conocimiento y consentimiento del usuario. La Figura 4 muestra la distribución del uso de permisos personalizados por aplicaciones que incorporan TPL. Encontramos que los permisos más utilizados - `READ_LOGS` — que permite que la aplicación (y, por lo tanto, los TPL dentro de ella) lean registros del sistema. No hay diferencias significativas entre los tres tipos de TPL de interés. Para completar, también encontramos que 94 aplicaciones con TPLs de interés también solicitan permisos personalizados. Curiosamente, el 53 % de los 88 permisos personalizados utilizados por estas aplicaciones son definidos y expuestos por Samsung.

### *IV-D. Exposición de componentes*

Los permisos personalizados no son el único mecanismo disponible para que los desarrolladores de aplicaciones expongan (o accedan) a componentes de otras aplicaciones. Las aplicaciones de Android también pueden interactuar entre sí mediante “Intents”, una abstracción de comunicación de alto nivel [41]. Una aplicación puede exponer

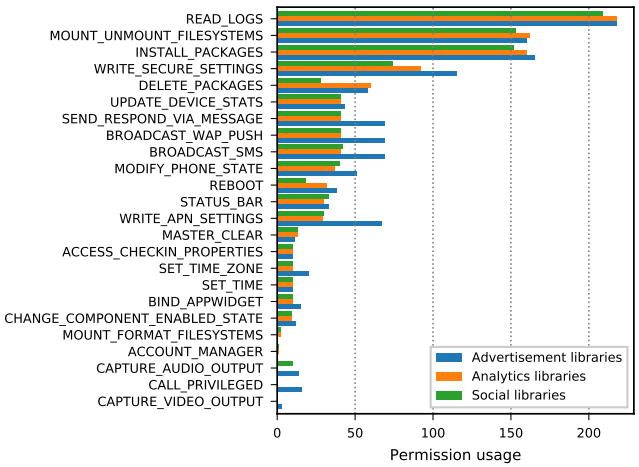


Figura 4: Permisos de sistema solicitaods por apps preinstaladas que incluyen librerías de terceros (TPLs).

sus componentes a aplicaciones externas declarando `android:exported=true` en el manifiesto sin proteger el componente con cualquier medida adicional, o agregando uno o más filtros a su declaración en el manifiesto; exponiéndolo a un tipo de ataque conocido en la literatura [78]. Si se usa el atributo `exported`, se puede proteger agregando un permiso al componente, ya sea un permiso personalizado o uno declarado por AOSP, a través de la verificación de forma programática de los permisos de la aplicación de llamada.

Intentamos identificar prácticas de desarrollo potencialmente peligrosas y pobres que pueden conducir a que los componentes sean expuestos sin ningún tipo de protección adicional. Exta exportación de componentes puede conducir a: $i)$ aplicaciones dañinas o maliciosas que inician una actividad expuesta, engañando así a los usuarios que creen que están interactuando con el software benigno; $ii)$ aplicaciones que pueden iniciar y vincularse a servicios desprotegidos; y $iii)$ aplicaciones maliciosas que obtienen acceso a datos confidenciales o la capacidad de modificar el estado interno de la aplicación.

Encontramos 6.849 aplicaciones preinstaladas que potencialmente exponen al menos una actividad en dispositivos de 166 proveedores y firmadas por 261 firmas de desarrollador con `exported=true`. Para los servicios, 4,591 aplicaciones (presentes en 157 proveedores) firmadas por 183 desarrolladores, incluidos los fabricantes, potencialmente exponen uno o más de sus servicios a aplicaciones externas. Los 10 principales fabricantes en nuestro dataset representan más del 70 % de las actividades y servicios potencialmente expuestos. Otros ejemplos relevantes incluyen una aplicación que parece exponer varias actividades relacionadas con las configuraciones del sistema (administración de dispositivos, redes, etc.), lo que podría permitir a un desarrollador externo acceder o incluso alterar la configuración del dispositivo. El paquete principal `com.android.mms` que ha sido personalizado en las versiones de Android desarrolladas por varios fabricantes también exponen los servicios para leer mensajes WAP a otras aplicaciones. También encontramos 8 diferentes instancias de una aplicación desarrollada por un tercero de la cadena de ensamblaje y que se encuentra en teléfonos de dos grandes fabricantes de Android, cuyo propósito es proporcionar soporte técnico remoto a los clientes. Este servicio en particular proporciona administración remota del dispositivo al MNO, incluido el capacidad de grabar audio y video, buscar archivos, acceder a la configuración del sistema y cargar / descargar archivos. El servicio clave para hacerlo es expuesto y puede ser abusado por otras aplicaciones.

Dejamos el estudio detallado de las aplicaciones vulnerables a este ataque y el estudio del acceso a estos recursos por aplicaciones disponibles públicamente en Google Play para trabajos futuros.

## V. Análisis del Comportamiento de Apps Preinstaladas

En esta sección analizamos la presencia de comportamientos potencialmente dañinos y no deseados. Para esto, aprovechamos herramientas disponibles en la literatura de análisis estático y dinámico de software Android y así caracterizar el propósito de dichas prácticas en las apps preinstaladas que se han obtenido, incluyendo el acceso invasivo a datos sensibles y privados del usuario.

### *V-A. Análisis estático*

En esta sección, seleccionamos todas las aplicaciones preinstaladas para determinar la presencia de comportamientos potencialmente dañinos. Este paso nos permite obtener una visión general de los comportamientos en todo el conjunto de datos y también nos proporciona la base para calificar aplicaciones y marcar aquellas más interesantes que requieran un análisis más exhaustivo. Este paso es crítico ya que solo podríamos permitirnos

inspeccionar manualmente un subconjunto limitado de todas las aplicaciones disponibles.

**Herramientas de Análisis** Nuestro método de estudio integra varias herramientas de análisis estático específico para aplicaciones de Android, incluido Androwarn [11], FlowDroid [73], y Amandroid [91], así como una serie de scripts personalizados basados en Apktool [12] y Androguard [3]. En esta etapa no utilizamos herramientas de análisis dinámico, lo que nos impide identificar comportamientos ocultos que se basen en la carga dinámica de código (carga DEX) o "reflection". Esto significa que nuestros resultados presentan una estimación a la baja de todos los posibles comportamientos potencialmente dañinos. Buscamos, no obstante, aplicaciones que carguen código dinámicamente y usen técnicas de "reflection" para identificar objetivos que merezcan una inspección manual.

**Dataset.** Debido a la escala de nuestro dataset — el cual consta de 82.501 archivos APK con 6.496 nombres de paquetes únicos — seleccionamos aleatoriamente un archivo APK para cada nombre de paquete y analizamos el conjunto resultante de aplicaciones, obteniendo un informe de análisis para 48 % de ellos. La mayoría de los paquetes restantes no se pudieron analizar debido a la ausencia de un `classes.dex` para archivos indexados. Aunque en algunos casos teníamos el archivo `.odex` correspondiente, generalmente no podíamos de-odexizarlos ya que se necesitaba el archivo de marco de Android del dispositivo para completar este paso. Además, no pudimos analizar un pequeño subconjunto de aplicaciones debido a las limitaciones de nuestras herramientas, incluidos los errores generados durante el análisis, el tamaño del archivo, o limitacions inherentes a las herramientas de análisis que dejan de responder después de horas de procesamiento. Por tanto, centramos nuestro análisis en el subconjunto de aplicaciones para las cuales pudimos generar informes satisfactoriamente.

**Resultados.** El estudio de los informes de análisis proporcionados por nuestras herramientas nos ha permitido identificar la presencia de 36 comportamientos potencialmente intrusivos de privacidad, o comportamientos potencialmente dañinos. Estos son enumerados en la Tabla VI. Los resultados sugieren que una fracción significativa de las aplicaciones analizadas podría acceder y diseminar datos sensibles del usuario y del dispositivo, incluyendo la ubicación del usuario y la configuración actual del dispositivo. De acuerdo a nuestro análisis, estos resultados dan la impresión de que la recopilación y difusión de datos personales (independientemente del propósito o consentimiento) no solo es generalizado en la industria móvil sino que también viene preinstalado. Otros comportamientos concernientes a priori incluyen la posible diseminación de contactos y contenido de SMS (164 y 74 aplicaciones, respectivamente), envío de SMS (29 aplicaciones) y los registros de llamadas telefónicas (339 aplicaciones). Aunque existen casos de uso perfectamente legítimos para estos comportamientos, también prevalecen en casos nocivos y potencialmente no deseados. La distribución del número de comportamientos potencialmente dañinos por aplicación sigue una distribución potencial. El 25 % de las aplicaciones analizadas presentan al menos 5 de estos comportamientos, y casi el 1 % de las aplicaciones muestran 20 o más. La mayor parte de la distribución se relaciona con la recopilación de identificadores de usuario, red y telefonía, e incluso información obtenida del Package Manager de Android. Esto proporciona una idea de cuán generalizadas son las huellas digitales de usuarios y dispositivos en la actualidad.

### *V-B. Análisis de tráfico*

Si bien el análisis estático puede ser útil para determinar un límite inferior de lo que es capaz de hacer una aplicación, esta técnica da una imagen incompleta del comportamiento real de una aplicación. Esto podría deberse a rutas de código huerfanas que nunca llegan a ejecutarse, incluidos los que se encuentran dentro de bibliotecas vinculadas estática y dinámicamente que no se proporcionan con aplicaciones, e incluso comportamientos determinados por la lógica del lado del servidor (p. Ej., debido a ofertas de anuncios en tiempo real) o código que se carga en tiempo de ejecución utilizando las técnicas de "reflection". Esta limitación de los enfoques estáticos generalmente se aborda complementando el análisis estático con el uso de herramientas de análisis dinámico con las que obtener evidencia real de comportamientos. Sin embargo, debido a varias limitaciones (incluidas las dependencias con hardware y componentes de software específicos al dispositivo) no es factible en nuestro caso la ejecución de todas las aplicaciones preinstaladas en nuestro dataset de forma dinámica. En cambio, decidimos usar el conjunto de datos de tráfico móvil de Lumen para encontrar evidencia de la difusión de datos personales de

| Tipo / comportamientos de PII accedidos | | Apps (#) | Apps (%) |
|---|---|---|---|
| Identificadores de telefonía | IMEI | 687 | 21.8 |
| | IMSI | 379 | 12 |
| | Número de teléfono | 303 | 9.6 |
| | MCC | 552 | 17.5 |
| | MNC | 552 | 17.5 |
| | Nombre del operador | 315 | 10 |
| | Numero de serie SIM | 181 | 5.7 |
| | Estado SIM | 383 | 12.1 |
| | País actual | 194 | 6.2 |
| | País SIM | 196 | 6.2 |
| | Número de voicemail | 29 | 0.9 |
| Configuración del dispositivo | Versión del software | 25 | 0.8 |
| | Estado del teléfono | 265 | 8.4 |
| | Apps instaladas | 1,286 | 40.8 |
| | Tipo de teléfono | 375 | 11.9 |
| | Logs | 2,568 | 81.4 |
| Ubicación | GPS | 54 | 1.7 |
| | Ubicación de la celda | 158 | 5 |
| | CID | 162 | 5.1 |
| | LACA | 137 | 4.3 |
| Interfaces de red | Configuración de Wi-Fi | 9 | 0.3 |
| | Red actual | 1,373 | 43.5 |
| | Plan de datos | 699 | 22.2 |
| | Estado de conexión | 71 | 2.3 |
| | Tipo de red | 345 | 10.9 |
| Información personal | Contactos | 164 | 11 |
| | SMS | 73 | 2.31 |
| Uso abusivo de servicios de telefonía | Envío de SMS | 29 | 0.92 |
| | Intercepción de SMS | 0 | 0 |
| | Desactivar SMS notif. | 0 | 0 |
| | Llamadas telefónicas | 339 | 10.7 |
| Interceptación de audio/video | Grabación de audio | 74 | 2.4 |
| | Captura de video | 21 | 0.7 |
| Ejecución de código arbitrario | Código nativo | 775 | 24.6 |
| | Comandos de Linux | 563 | 17.9 |
| Conexión remota | Conexión remota | 89 | 2.8 |

Cuadro VI: Número de apps preinstaladas que acceden a datos sensibles o muestran comportamientos potencialmente peligrosos. El porcentaje se refiere a la subconjunto de paquetes triaged ($N = 3,154$).

las aplicaciones preinstaladas presentes en ambos datasets.

**Resultados.** De las 3.118 aplicaciones preinstaladas con permisos de acceso a Internet, 1.055 han sido observadas por Lumen en dispositivos reales a través de técnicas de crowdsourcing. En este punto, nuestro análisis de aplicaciones preinstaladas se centró en dos aspectos principales: descubrir el ecosistema de las organizaciones que poseen los dominios a los que se conectan estas aplicaciones, y analizar los tipos de información privada que los dispositivos de los usuarios diseminan a estas organizaciones. Para comprender el ecosistema de la recopilación de datos mediante aplicaciones preinstaladas, estudiamos dónde se recopilan los datos de estas aplicaciones y donde hace su primera parada. Usamos los nombres de dominios completos (FQDN) de los servidores con los que contactan las apps presintaladas y usamos técnicas de minería de texto descritas en nuestro trabajo anterior [84] para determinar la organización matriz tras estos dominios.

**Los grandes actores.** La Tabla VII muestra las organizaciones principales que poseen los dominios más populares contactados por aplicaciones preinstaladas presentes en el dataset de Lumen. De los 54,614 dominios contactados por estas aplicaciones, 7,629 pertenecen a servicios de publicidad y tracking en línea, conocidos como servicios ATS [84]. Estos servicios están representados por organizaciones como Alphabet, Facebook, Verizon (ahora propietaria de Yahoo!, AOL y Flurry), Twitter (organización matriz de MoPub), AppsFlyer, comScore y otros más. Como se esperaba, Alphabet, la entidad que posee y mantiene la plataforma Android y algunos de los mayores servicios de publicidad y seguimiento en línea (ATS) [84], también posee la mayoría de los dominios a los que se conectan las aplicaciones preinstaladas. Además, los vendedores que incorporan en sus dispositivos los servicios básicos de Google para Android como la Google Play Store tiene que pasar por el programa de certificación de Google que, en parte, implica precargar los dichos servicios. Entre estos servicios se encuentra el paquete `com.google.backuptransport` de Google, que envía una variedad de información sobre el usuario y el dispositivo en el que se ejecuta a los servidores de Google.

El análisis de tráfico también confirma que los servicios de Facebook y Twitter vienen preinstalados en muchos teléfonos y están integrados en varias aplicaciones. Muchos dispositivos también preinstalan aplicaciones meteorológicas como AccuWeather y The Weather Channel. De acuerdo a los hallazgos de varios esfuerzos de investigación previos, estos proveedores de contenido metereológico también recopilan información sobre los dispositivos y sus usuarios [86], [84].

### *V-C. Análisis manual: casos relevantes*

Estudiamos la salida proporcionada por nuestro metodología de análisis estático y dinámico para clasificar aplicaciones y así marcar un subconjunto reducido de paquetes que inspeccionar manualmente. Nuestro objetivo aquí era identificar con precisión el comportamiento potencialmente dañino y no deseado en aplicaciones preinstaladas. Se

| Organización | # de apps | de dominios |
|---|---|---|
| Alphabet | 566 | 17052 |
| Facebook | 322 | 3325 |
| Amazon | 201 | 991 |
| Verizon Communications | 171 | 320 |
| Twitter | 137 | 101 |
| Microsoft | 136 | 408 |
| Adobe | 116 | 302 |
| AppsFlyer | 98 | 10 |
| comScore | 86 | 8 |
| AccuWeather | 86 | 15 |
| MoatInc. | 79 | 20 |
| Appnexus | 79 | 35 |
| Baidu | 72 | 69 |
| Criteo | 70 | 62 |
| PerfectPrivacy | 68 | 28 |
| Other ATS | 221 | 362 |

Cuadro VII: Las 15 principales organizaciones ATS por número de aplicaciones que se conectan a todos sus dominios asociados.

agregaron otras aplicaciones a este conjunto en función de los resultados de nuestra biblioteca de terceros y el análisis de permisos realizado en las secciones III y IV, respectivamente. Analizamos manualmente 158 aplicaciones usando herramientas estándar que incluyen desensambladores DEX (baksmali), descompiladores DEX a java (jadx, dex2jar), herramientas de análisis de recursos (Apktool), herramientas de instrumentación (Frida) y métodos de ingenieria inversa (radare2 e IDA Pro) para el análisis de código nativo. Nuestros principales hallazgos se pueden agrupar en tres grandes categorías: 1) malware conocido; 2) software con posible acceso y difusión de datos personales; y 3) aplicaciones potencialmente dañinas. La Tabla VIII proporciona algunos ejemplos para cada tipo de comportamientos que encontramos.

**Malware conocido.** Nos encontramos con varias instancias aisladas de malware conocido en la partición del sistema, principalmente en dispositivos de gama baja pero también en algunos teléfonos de gama alta. Identificamos variantes de familias conocidas de malware de Android que han sido prevalentes en los últimos años, incluyendo Triada, Rootnik, SnowFox, Xinyin, Ztorg, Iop y software dudoso desarrollado por GMobi. Utilizamos VirusTotal para etiquetar estas muestras. Según los informes AV existentes, el rango de comportamientos que tales muestras abarca el fraude bancario, el envío de SMS a números premium o la suscripción a servicios de pago, la instalación silenciosa de aplicaciones, o fraude publicitario, entre otros. Si bien nuestro método no nos permite distinguir si potencialmente las aplicaciones maliciosas están preinstaladas o aprovecharon las vulnerabilidades del sistema para instalarse en la partición del sistema, es importante destacar que la presencia de malware preinstalado en dispositivos Android ha sido reportada previamente por varias fuentes[65], [5], [66]. Algunas de las muestras encontradas utilizan servidores de Comando y Control (C2) que todavía están en funcionamiento en el momento de escribir este documento.

**Acceso y difusión de datos personales.** Casi todas las aplicaciones que identificamos como capaces de acceder a PII, parecen difundirlo a servidores de terceros. También observamos instancias de aplicaciones con capacidad para obtener huellas (fingrepritns) de los dispositivos a través del hardware y redes, a menudo recopiladas bajo el término “capacidad del dispositivo”, e incluso servicios de análisis que rastrean la instalación y eliminación de aplicaciones (en particular, aplicaciones de noticias, como las realizadas por CNBC, Bloomberg, TechCrunch y The Economist, entre otras). Los comportamientos más intrusivos abarcan aplicaciones capaces de recopilar y enviar correos electrónicos y metadatos de registros de llamadas telefónicas. El caso más extremo que nosotros hemos analizado es un servicio de recopilación de datos contenido en un servicio FOTA asociado con Redstone Sunshine Technology Co., Ltd. [60], un proveedor de OTA que admite tener acceso a “550 millones de usuarios de teléfonos y socios de IoT en 40 países” [21]. Esta aplicación incluye un servicio que puede recopilar y difundir docenas de elementos de datos, incluidos identificadores de usuario y dispositivo, información de comportamiento (recuentos de SMS y llamadas enviadas y recibidas, y estadísticas sobre flujos de red) y estadísticas de uso e información de rendimiento por paquete instalado. En general, este software parece implementar un programa de análisis que admite varias estrategias de monetización, desde la orientación optimizada de anuncios hasta proporcionar datos de rendimiento tanto a los desarrolladores como a los fabricantes. Hacemos hincapié en que estos datos recopilados no solo son notablemente extensos y multidimensionales, sino que también están muy lejos de ser anónimos ya que son vinculado con IDs de usuario y dispositivo.

**Aplicaciones potencialmente peligrosas.** Encontramos 612 aplicaciones preinstaladas que potencialmente implementan funciones avanzadas de ingeniería de acuerdo con sus paquetes y nombres de aplicaciones. Dichas funciones incluyen tareas

| Familia | Comportamiento potencial y prevalencia |
|---|---|
| **Malware conocido** | |
| Triada | Disemina PII y otros datos confidenciales (SMS, registros de llamadas, datos de contacto, imágenes y videos almacenados). Descargas etapas adicionales. Rootea el dispositivo para instalar aplicaciones adicionales. |
| Rootnik [61] | Obtiene acceso root al dispositivo. Disemina PII e instala aplicaciones adicionales. Utiliza técnicas anti-análisis y anti-depuración. |
| GMobi [10], [66] | Servicio de Comercio Gmobi. Diseminación de PII, incluido el número de serie del dispositivo y la dirección MAC, la geolocalización, los paquetes instalados y los correos electrónicos. Recibe comandos de servidores remotos para (1) enviar un SMS a un número dado; (2) descargar e instalar aplicaciones; (3) visite un enlace; o (4) muestra una ventana emergente. Esoha sido identificado en dispositivos de gama baja. |
| **Aplicaciones potencialmente peligrosas** | |
| Aplicación de rooteo | Expone un receptor desprotegido que arranca el dispositivo al recibir un código secreto de telefonía (mediante intento o marcando `*#*#9527#*#*`). |
| Bloqueador | Si el dispositivo no contiene un archivo firmado en una ubicación en particular, carga y aplica 2 listas negras: una que contiene 103 paquetes asociados con aplicaciones de evaluación comparativa, y otra con 56 dominios web relacionados con tecnologías telefónicas. |
| **Posible acceso y difusión de datos personales** | |
| TrueCaller | Envía PII a sus propios servidores y también a ATSes de terceros integrados en la app, como AppsFlyer, MoPub, propiedad de Twitter, Crashlytics, inMobi, Facebook,y otros. Carga datos de llamadas telefónicas a al menos uno de sus propios dominios. |
| MetroName ID | Difunde PII a sus propios servidores y también a servicios de terceros embedidos como Piano, un servicio para la medición de audiencia de medios y de análisis de rastreo de la instalación de aplicaciones de noticias y otros socios, incluidos CNBC, Bloomberg, TechCrunch y The Economist, entre otros, la presencia de los cuales informa a sus propios dominios. |
| Adups [46] | Aplicación FOTA. Recopila y comparte PII privado y privado con sus propios servidores y los de dominios ATS de terceros integrados, incluido Advmob y Nexage. Encontrado en todo el mundo en 55 marcas diferentes. |
| Stats/Meteor | Servicio FOTA de Redstone. Utiliza la carga de código dinámico y técnicas de “reflection” para implementar componentes ubicados en 2 archivos DEX cifrados. Disemina más de 50 tipos de datos que caracterizan completamente el hardware, el servicio de telefonía, la red, la geolocalización y los paquetes instalados. Realiza perfiles de comportamiento y rendimiento, incluidos recuentos de SMS / MMS, registros de llamadas, bytes enviados y transmitidos, y estadísticas de uso y rendimiento contadores por paquete. Instala silenciosamente paquetes en el dispositivo e informa qué paquetes son instalados / eliminados por el usuario. |

Cuadro VIII: Ejemplos de casos relevantes y sus posibles comportamientos encontrados después del análisis manual sobre un subconjunto de aplicaciones de interés. Cuando se referencia a la difusión de datos personales, el término PII abarca los elementos enumerados en la Tabla VI.

relativamente inofensivas, como pruebas de hardware, pero también funciones potencialmente peligrosas como la capacidad de rootear el dispositivo. Encontramos instancias de tales aplicaciones en las que la función de rooting no estaba protegida en su manifiesto (es decir, el componente estaba disponible para que lo utilizara cualquier otra aplicación). También identificamos aplicaciones de modo de ingeniería vulnerable bien conocidas como MTKLogger [81]. Dichas aplicaciones exponen componentes desprotegidos que pueden ser mal utilizado por otras aplicaciones ubicadas en el dispositivo. Otros ejemplos incluyen un servicio de fabricante bien conocido, que bajo ciertas condiciones ponen en una lista negra las conexiones a una lista predefinida de 56 dominios web (revisión de dispositivos móviles y evaluación comparativa sitios web, principalmente) y desactiva cualquier paquete instalado que coincida con una de una lista de 103 aplicaciones de evaluación comparativa.

## VI. Limitaciones del estudio

**Completud y cobertura.** Nuestro conjunto de datos no es completo en términos de proveedores y modelos de Android, a pesar de que cubrimos aquellos con una mayor penetración de mercado, tanto de gamas altas y bajas. Nuestro proceso de recopilación de datos también se realiza con el mejor esfuerzo. La falta de conocimientos previos sobre la problemática de apps preinstaladas y la falta de documentación requería realizar un estudio detallado caso por caso y un cantidad significativa de inspección manual. En términos de aplicaciones analizadas, determinar la cobertura de nuestro estudio es difícil ya que no sabemos el número total de aplicaciones preinstaladas en todos los teléfonos analizados.

**Atribución.** Actualmente, no hay una manera fiable de identificar con precisión el desarrollador legítimo de una aplicación preinstalada dado que estas son autofirmadas. Hemos encontrado instancias de cer-

tificados con solo un código de país en el campo emisor, y otros con cadenas que sugieren que los principales proveedores (por ejemplo, Google) firmaron la aplicación, donde las aplicaciones ciertamente no fueron firmadas por ellos. Lo mismo se aplica a los nombres de paquetes y permisos, muchos de los cuales son opacos y su nomenclatura no sigue las recomendaciones de nomenclatura indicadas por Google. Del mismo modo, la falta de documentación sobre la definición y propósito de los permisos personalizados nos impidió automatizar nuestro análisis. Un estudio más profundo de este problema requeriría verificar si esos permisos se otorgan en tiempo de ejecución, rastreando el código para identificar completamente sus propósito y averiguar si otras aplicaciones los utilizan.

**Package Manager.** No obtenemos muestras del archivo packages.xml de los dispositivos de nuestros usuarios ya que puede contener información sensible sobre todos paquetes instalados, y no solo preinstalados. Consideramos que recopilar este archivo sería invasivo. Esto, sin embargo, limita nuestra capacidad de ver si las aplicaciones instaladas por el usuario están utilizando servicios expuestos por aplicaciones preinstaladas a través de “intents” o permisos personalizados. Tratamos de compensar esta falta con una búsqueda de aplicaciones públicas que usan permisos personalizados preinstalados, como se discute en Sección IV-D.

**Cobertura de comportamiento.** Nuestro estudio se centra en el análisis estático de las muestras recolectadas a través de Firmware Scanner, y sólo se aplicó el análisis dinámico a un subconjunto seleccionado de 1.055 paquetes. Esto nos impide analizar comportamientos sospechosos como la carga dinámica de código en tiempo de ejecución, o el uso de técnicas de “reflection”. A pesar de esto, nuestra línea de análisis sirvió para identificar una cantidad considerable de comportamientos potencialmente dañinos en software preinstalado. Un estudio más profundo y amplio posiblemente podría revelaría más casos.

**Identificación de dispositivos rooteados.** No hay una forma segura de saber programáticamente si un dispositivo está rooteado o no. Mientras nuestro conservador limita el número de falsos negativos, hemos encontrado casos de dispositivos con ROM alternativas conocidas que fueron no marcados como rooteados por RootBeer. Además, hemos encontrado algunas aplicaciones que permiten a un tercero rootear el dispositivo sobre la marcha para, por ejemplo, instalar nuevas aplicaciones en la partición del sistema como se discutió en la Sección VC. Algunas de estas aplicaciones pueden potencialmente “desrootear” el teléfono para evitar la detección. Ante la presencia de una aplicación de este tipo en un dispositivo, no podemos saber con certeza si un paquete determinado — particularmente una aplicación potencialmente maliciosa — fue instalada previamente por un actor en la cadena de suministro o si se instaló después.

## VII. Trabajos relacionados

**Personalización de imágenes de Android.** Algunos trabajos previos han estudiado las modificaciones realizadas en las imágenes AOSP, ya sea agregando certificados raíz [88], personalizando las aplicaciones predeterminadas [72] o el sistema operativo [94]. Aafer [71] introdujo una nueva clase de vulnerabilidad causada por el proceso de personalización del firmware. Si se elimina una aplicación pero queda una referencia a ella en el sistema operativo, una aplicación maliciosa podría suplantarla, lo que podría generar problemas de privacidad y seguridad. Si bien estos estudios se han centrado en las imágenes de Android en su conjunto, en lugar de las aplicaciones preinstaladas, todas muestran la complejidad del ecosistema de Android y subrayan la falta de control sobre la cadena de suministro.

**Permisos de Android.** Estudios anteriores sobre permisos de Android han usado principalmente las técnicas de análisis estático para inferir el papel de un permiso dado [74], [77]. Estos estudios, sin embargo, no cubren las versiones más recientes de Android [93] ni permisos proprietarios o personalizados. En [80], Jiang demostró cómo se usan los permisos personalizados para exponer y proteger los servicios. Nuestro trabajo complementa este estudio al mostrar cómo los fabricantes de dispositivos y terceros declaran y usan permisos personalizados, dando el primer paso hacia un análisis completo y en profundidad de todo el panorama de los permisos personalizados.

**Vulnerabilidades en aplicaciones preinstaladas.** Un artículo reciente de Wu et al. [92] también utilizó mecanismos de crowdsourcing para detectar aplicaciones que escuchan un determinado puerto TCP o UDP y analizan las vulnerabilidades causadas por esta práctica. Mientras su estudio no se limita a las aplicaciones instaladas por el usuario, muestran evidencia de aplicaciones preinstaladas que exhiben este comportamiento

## VIII. Discusión and Conclusiones

Este artículo estudió, a escala, el vasto e inexplorado ecosistema de software preinstalado en dispositivos Android y su impacto potencial en la privacidad y derechos de los consumidores. Este estudio ha dejado en claro que, gracias en gran parte a la naturaleza de código abierto de la plataforma Android y la complejidad de su cadena de suministro, organizaciones de diversos tipos y tamaños tienen la capacidad de introducir su software en el firmware de Android. Como demostramos en este estudio, esta situación se ha convertido en un peligro para la privacidad de los usuarios e incluso seguridad debido a un abuso de privilegios o como resultado de malas prácticas de ingeniería de software que introducen vulnerabilidades y puertas traseras peligrosas.

**La cadena de suministros.** La gran cantidad de actores involucrados en el desarrollo de software preinstalado y la gama de la cadena de suministro desde fabricantes de hardware hasta MNOs y servicios de publicidad y tracking. Estos actores tienen acceso privilegiado a los recursos del sistema a través de su presencia en aplicaciones preinstaladas, pero también como bibliotecas de terceros incrustadas en ellas. Potenciales asociaciones y acuerdos comerciales, realizados a puerta cerrada entre las partes interesadas, han convertido los datos de los usuarios en una mercancía antes que los usuarios comprar sus dispositivos o decidir instalar su propio software.

**Atribución.** Desafortunadamente, debido a la falta de una autoridad certificadora central o de un sistema de confianza para permitir la verificación y la atribución de certificados firmados que se utilizan para firmar aplicaciones preinstaladas, y debido a la falta de cualquier mecanismo para identificar el propósito y la legitimidad de muchas de estas aplicaciones y permisos personalizados, es difícil atribuir comportamientos no deseados y perjudiciales a la parte o partes responsables Esto tiene implicaciones negativas más amplias para la responsabilidad de los agentes que forman este ecosistema y la cadena de suministros en su conjunto.

**El papel de los usuarios y el consentimiento informado.** Mientras tanto, los usuarios de dispositivos Android no son, en general, conscientes de la presencia de la mayoría del software que viene preinstalado en sus dispositivos ni de los riesgos que estas suponen a su privacidad. Los usuarios no tienen idea de las diversas relaciones y asociaciones que existen entre agentes para el intercambio de datos personales, ni tienen capacidad para decidir qué viene preinstalado en sus teléfonos. Las actividades, los datos personales y los hábitos de los usuarios pueden ser trackeados constantemente por partes interesadas y agentes de los que muchos de ellos nunca han oído hablar, y mucho menos han consentido en recopilar sus datos. Hemos demostrado instancias de dispositivos con software desarrollado por compañías con la capacidad de rootear y controlar dispositivos de forma remota sin el conocimiento del usuario, y con la capacidad de instalar aplicaciones a través de campañas de monetización y adquisición de usuarios. Incluso si los usuarios decidiesen detener o eliminar algunas de estas aplicaciones, no podrá hacerlo, ya que muchos de ellos son servicios básicos de Android y otros no pueden ser eliminados permanentemente por el usuarios sin privilegios de root. Tampoco está claro si los usuarios realmente han dado su consentimiento a estas prácticas, o si fueron informados sobre ellas antes de usar los dispositivos (es decir, en el primer arranque) en primer lugar. Para aclarar esta incógnita, adquirimos 6 dispositivos Android populares completamente nuevos de proveedores como Nokia, Sony, LG y Huawei a través de un gran minorista español. Al arrancarlos, 3 dispositivos no presentan una política de privacidad, solo los términos de servicio de Android. El resto presentó una política de privacidad que solo menciona que recopilan datos sobre el usuario, incluidos datos sensibles como el IMEI para ofrecer servicios de valor agregado. Tenga en cuenta que los usuarios no tienen otra opción salvo aceptar los términos de servicio de Android, así como los del fabricante. De lo contrario, Android simplemente dejará de arrancar haciendo que el dispositivo sea inutilizable.

**Reglamento de Protección al Consumidor.** Mientras que algunas jurisdicciones tienen regulaciones muy laxas que rigen el seguimiento y los datos en línea recopilación, ha habido una serie de movimientos para regular y controlar estas prácticas de forma más estricta, como el GDPR en la UE [28],y CCPA de California [20] en los Estados Unidos. Si bien estos esfuerzos son ciertamente útiles para regular la invasión desenfrenada en la privacidad de los usuarios, éstas aún tienen un largo camino por recorrer. La mayoría de los dispositivos móviles aún carecen de un mecanismo claro y significativo para obtener el consentimiento informado, lo cual es una posible violación del GDPR. De hecho, es posible que muchos de los ATS embedidos en el software

preinstalado en dispositivos Android puede no ser compatible con COPPA [87] - una regla federal de EE. UU. para proteger a los menores de ilegal seguimiento en línea [22] -, a pesar del hecho de que muchos menores en los Estados Unidos usan dispositivos móviles. Esto indica que incluso en jurisdicciones con leyes y regulaciones estrictas para preservar la privacidad y protección del consumidor, todavía queda una gran brecha entre lo que se hace en la práctica y las capacidades de aplicación de las agencias designadas para hacer cumplir la ley.

**Recomendaciones** Para limitar el impacto de los problemas mencionados anteriormente y hacer que el ecosistema sea más transparente, proponemos una serie de recomendaciones que se realizan bajo el supuesto de que los interesados están dispuestos a autorregularse y mejorar el status quo. Somos conscientes de que algunas de estas sugerencias pueden inevitablemente no alinearse con los intereses corporativos de todas las organizaciones en la cadena de suministro, y que puede ser necesario la presencia de terceros para auditar y certificar el proceso independientemente. Google podría ser un candidato principal para ello dada su capacidad para otorgar licencias a los vendedores y sus programas de certificación. Alternativamente, en ausencia de auto-regulación, los gobiernos y los organismos reguladores podrían intervenir y promulgar regulaciones y ejecutar acciones que restituyan parte del control de los diversos actores en la cadena de suministro. También proponemos una serie de acciones que ayudarían a los investigadores independientes para detectar comportamientos engañosos y potencialmente dañinos por parte del software preinstalado.

• *Atribución y responsabilidad* para combatir la dificultad en la atribución y la consiguiente falta de responsabilidad, proponemos la introducción y uso de certificados firmados por autoridades de certificación de confianza global. Alternativamente, puede ser posible para construir un repositorio de transparencia de certificados dedicado a proporcionar detalles y atribución de certificados auto firmados utilizados para firmar varias aplicaciones de Android, incluidas las preinstaladas.

• *Documentación accesible y formularios de consentimiento* similar a la forma en que los componentes de código abierto de Android requieren cualquier versión modificada del código que se pondrá a disposición del público, se puede requerir que los dispositivos Android documenten el conjunto específico de aplicaciones que se han preinstalado, junto con su propósito y la entidad responsable de cada pieza de software, de manera que sea accesible y comprensible para los usuarios finales. Esto asegurará que exista al menos un punto de referencia para que los usuarios (y los reguladores) puedan encuentrar información precisa sobre las aplicaciones preinstaladas y sus prácticas o fines. Además, los resultados de nuestro análisis de la presencia de formularios de consentimiento a pequeña escala en algunos fabricantes de Android deja mucho que desear desde una perspectiva de transparencia: los usuarios no están suficientemente informados sobre el software de terceros que está preinstalado en sus dispositivos, incluidos los servicios integrados de tracking y publicidad en línea de terceros, el tipos de datos que recopilan de ellos de forma predeterminada, y las asociaciones que permiten compartir datos personales a través de Internet. Esto requiere una nueva forma de política de privacidad adecuada al contexto de las aplicaciones preinstaladas (y se apliquen) para garantizar que las prácticas se comunican al menos al usuario de manera clara y accesible. Esto debería ir acompañado de mecanismos para permitir a los usuarios tomar decisiones informadas sobre cómo, o si usar, dichos dispositivos sin tener que rootearlos.

**Observaciones finales.** A pesar de que este estudio ha requerido más de un año de esfuerzos, solo pudimos arañar la superficie de un problema mucho mayor. Este trabajo es, por lo tanto, exploratorio, y esperamos que atraiga más atención sobre el ecosistema de la cadena de suministro de Android y su impacto en privacidad y seguridad de los usuarios. Hemos discutido nuestros resultados con Google, que nos dio comentarios útiles. Nuestro trabajo también fue la base de un informe elaborado por la Agencia Española de Protección de Datos (AEPD) [3]. También mejoraremos las capacidades y características de Firmware Scanner y Lumen para abordar algunas de las limitaciones antes mencionadas y desarrollar métodos para realizar análisis dinámico de software preinstalado. Dada la escala del ecosistema y la necesidad de inspecciones manuales, nosotros gradualmente haremos que nuestro conjunto de datos (que sigue creciendo en el momento de escribir este artículo) esté disponible para la comunidad de investigación y reguladores para ayudar en futuras investigaciones y alentar más investigación en esta área.

## AGRADECIMIENTOS

Estamos profundamente agradecidos a nuestros usuarios de Firmware Scanner por permitir este estudio, y ElevenPaths por su apoyo inicial en este proyecto. Agradecemos a los revisores anónimos por sus útiles comentarios. Este proyecto ha sido parcialmente financiado por el US National Science Foundation (subvención CNS-1564329), el programa de acción de innovación Horizonte 2020 de la Unión Europea (acuerdo de subvención no. 786741, Proyecto SMOOTH), Ministerio de Ciencia, Innovación y Universidades de España (otorga DiscoEdge TIN2017-88749-R y SMOG-DEV TIN2016-79095-C2-2-R), y la Comunidad de Madrid (subvención EdgeData-CM P2018 / TCS-4499). Las opiniones, hallazgos, conclusiones o recomendaciones expresadas en este documento son de los autores y no reflejan las opiniones de los organismos de financiación.

## REFERENCIAS

[1] AdGuard - Meizu Incompatibilities. https://github.com/AdguardTeam/AdguardForAndroid/issues/800. [Online; accessed 31-March-2019].

[2] Amazon suspends sales of Blu phones for including preloaded spyware, again. https://www.theverge.com/2017/7/31/16072786/amazon-blu-suspended-android-spyware-user-data-theft. [Online; accessed 31-March-2019].

[3] Androguard. https://github.com/androguard/androguard/. [Online; accessed 31-March-2019].

[4] Android — Certified. https://www.android.com/certified/. [Online; accessed 31-March-2019].

[5] Android Adware and Ransomware Found Preinstalled on High-End Smartphones. https://www.bleepingcomputer.com/news/security/android-adware-and-ransomware-found-preinstalled-on-high-end-smartphones/. [Online; accessed 31-March-2019].

[6] Android Certified Partners. https://www.android.com/certified/partners/. [Online; accessed 31-March-2019].

[7] Android Compatibility Program Overview. https://source.android.com/compatibility/overview. [Online; accessed 31-March-2019].

[8] Android Developer Documentation. https://developer.android.com/. [Online; accessed 31-March-2019].

[9] Android Trackers. https://fiksu.com/resources/android_trackers/. [Online; accessed 31-March-2019].

[10] Android.Gmobi.1. https://vms.drweb.com/virus/?_is=1&i=7999623&lng=en. [Online; accessed 31-March-2019].

[11] Androwarn–Yet another static code analyzer for malicious Android applications. https://github.com/maaaaz/androwarn. [Online; accessed 31-March-2019].

[12] Apktool–A tool for reverse engineering Android apk files. https://ibotpeaches.github.io/Apktool/. [Online; accessed 31-March-2019].

[13] App Traps: How Cheap Smartphones Siphon User Data in DevelopingmCountries. https://www.wsj.com/articles/app-traps-how-cheap-smartphones-help-themselves-to-user-data-1530788404. [Online; accessed 31-March-2019].

[14] Application signing. https://developer.android.com/studio/publish/app-signing. [Online; accessed 31-March-2019].

[15] Appsee — Features. https://www.appsee.com/features. [Online; accessed 31-March-2019].

[16] Appsee Mobile App Analytics. https://www.appsee.com/. [Online; accessed 31-March-2019].

[17] Asurion. https://www.asurion.com/. [Online; accessed 31-March-2019].

[18] Baidu Geocoding API. http://api.map.baidu.com/lbsapi/geocoding-api.htm. [Online; accessed 31-March-2019].

[19] Baidu SDK. https://developer.baidu.com/. [Online; accessed 31-March-2019].

[20] California Consumer Privacy Act. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375. [Online; accessed 31-March-2019].

[21] China Mobile Network Partner Redstone Moves into Robotics. https://www.prweb.com/releases/2017/04/prweb14212503.htm. [Online; accessed 31-March-2019].

[22] COPPA - Children's Online Privacy Protection Act. http://coppa.org/. [Online; accessed 31-March-2019].

[23] CVE-2017-2709. https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2017-2709. [Online; accessed 31-March-2019].

[24] CVE-2017-2709. https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2015-0864. [Online; accessed 31-March-2019].

[25] Define a Custom Permission. https://developer.android.com/guide/topics/permissions/defining. [Online; accessed 31-March-2019].

[26] Digital Turbine - Privacy Policy. https://www.digitalturbine.com/privacy-policy/. [Online; accessed 31-March-2019].

[27] Estimote — indoor location with bluetooth beacons and mesh. https://estimote.com/. [Online; accessed 31-March-2019].

[28] EU General Data Protection Regulation (GDPR). https://eugdpr.org/. [Online; accessed 31-March-2019].

[29] Europe should be wary of Huawei, EU tech official says. https://www.reuters.com/article/us-eu-china-huawei-idUSKBN1O611X. [Online; accessed 31-March-2019].

[30] EXUS. https://www.exus.co.uk. [Online; accessed 31-March-2019].

[31] Facebook Gave Device Makers Deep Access to Data on Users and Friends. https://www.nytimes.com/interactive/2018/06/03/technology/facebook-device-partners-users-friends-data.html. [Online; accessed 31-March-2019].

[32] Facebook's Data Deals Are Under Criminal Investigation. https://www.nytimes.com/2019/03/13/technology/facebook-data-deals-investigation.html. [Online; accessed 31-March-2019].

[33] Firmware Scanner. https://play.google.com/store/apps/details?id=org.imdea.networks.iag.preinstalleduploader. [Online; accessed 31-March-2019].

[34] Global market share held by leading smartphone vendors. https://www.statista.com/statistics/271496/global-market-share-held-by-smartphone-vendors-since-4th-quarter-2009/. [Online; accessed 31-March-2019].

[35] GMobi — General Mobile Corporation. http://www.generalmobi.com/en/. [Online; accessed 31-March-2019].

[36] Google Cloud Messaging. https://developers.google.com/cloud-messaging/android/android-migrate-fcm. [Online; accessed 31-March-2019].

[37] Hiya. https://hiya.com/. [Online; accessed 31-March-2019].

[38] Hiya Partners. https://hiya.com/hiya-data-policy. [Online; accessed 31-March-2019].

[39] How does Truecaller get its data? https://support.truecaller.com/hc/en-us/articles/212638485-How-does-Truecaller-get-its-data. [Online; accessed 31-March-2019].

[40] Infinum Inc. https://infinum.co. [Online; accessed 31-March-2019].

[41] Intents and Intent Filters - Android Developers. https://developer.android.com/guide/components/intents-filters. [Online; accessed 31-March-2019].

[42] IronSource — App monetization done right. https://www.ironsrc.com/. [Online; accessed 31-March-2019].

[43] IronSource - AURA. https://company.ironsrc.com/enterprise-solutions/. [Online; accessed 31-March-2019].

[44] IronSource - Aura for Advertisers. https://www.slideshare.net/ironSource/aura-for-advertisers. [Online; accessed 31-March-2019].

[45] Kryptowire Discovers Mobile Phone Firmware that Transmitted Personally Identifiable Information (PII) without User Consent or Disclosure. https://www.kryptowire.com/adups_security_analysis.html. [Online; accessed 31-March-2019].

[46] Kryptowire Provides Technical Details on Black Hat 2017 Presentation: Observed ADUPS Data Collection & Data Transmission. https://www.kryptowire.com/observed_adups_data_collection_behavior.html. [Online; accessed 31-March-2019].

[47] locationlabs by Avast. https://www.locationlabs.com/. [Online; accessed 31-March-2019].

[48] Lumen Privacy Monitor. https://play.google.com/store/apps/details?id=edu.berkeley.icsi.haystack. [Online; accessed 31-March-2019].

[49] Manifest permissions. https://developer.android.com/reference/android/Manifest.permission. [Online; accessed 31-March-2019].

[50] Monetize, advertise and analyze Android apps. https://www.appbrain.com. [Online; accessed 31-March-2019].

[51] OnePlus Device Root Exploit: Backdoor in EngineerMode App for Diagnostics Mode. https://www.nowsecure.com/blog/2017/11/14/oneplus-device-root-exploit-backdoor-engineermode-app-diagnostics-mode/. [Online; accessed 31-March-2019].

[52] OnePlus left a backdoor in its devices capable of root access. http://www.androidpolice.com/2017/11/15/oneplus-left-backdoor-devices-capable-root-access/. [Online; accessed 31-March-2019].

[53] OnePlus OxygenOS built-in analytics. https://www.chrisdcmoore.co.uk/post/oneplus-analytics/. [Online; accessed 31-March-2019].

[54] OnePlus Secret Backdoor. https://www.theregister.co.uk/2017/11/14/oneplus_backdoor/. [Online; accessed 31-March-2019].

[55] Permissions overview. https://developer.android.com/guide/topics/permissions/overview.html. [Online; accessed 31-March-2019].

[56] Phone Number Search — TrueCaller. https://www.truecaller.com/. [Online; accessed 31-March-2019].

[57] Privacy Grade. http://privacygrade.org. [Online; accessed 31-March-2019].

[58] PrivacyStar. https://privacystar.com. [Online; accessed 31-March-2019].

[59] PrivacyStar Privacy Policy. https://privacystar.com/privacy-policy/. [Online; accessed 31-March-2019].

[60] Redstone. http://www.redstone.net.cn/. [Online; accessed 31-March-2019].

[61] Rootnik Android Trojan Abuses Commercial Rooting Tool and Steals Private Information. https://unit42.paloaltonetworks.com/rootnik-android-trojan-abuses-commercial-rooting-tool-and-steals-private-information/. [Online; accessed 31-March-2019].

[62] Simple to use root checking Android library. https://github.com/scottyab/rootbeer. [Online; accessed 31-March-2019].

[63] Smaato Blog. https://blog.smaato.com/everything-you-need-to-know-about-location-based-mobile-advertising. [Online; accessed 31-March-2019].

[64] Synchronoss Technologies - Privacy Policy. https://synchronoss.com/privacy-policy/#datacollected. [Online; accessed 31-March-2019].

[65] Triada Trojan Found in Firmware of Low-Cost Android Smartphones. https://www.bleepingcomputer.com/news/security/android-adware-and-ransomware-found-preinstalled-on-high-end-smartphones/. [Online; accessed 31-March-2019].

[66] Upstream - Low-end Android smartphones sold with pre-installed malicious software in emerging markets. https://www.upstreamsystems.com/pre-installed-malware-android-smartphones/. [Online; accessed 31-March-2019].

[67] VPN Service. https://developer.android.com/reference/android/net/VpnService. [Online; accessed 31-March-2019].

[68] What is “com,facebook,app manager” and why is it trying to download Instagram, Facebook, and Messenger. https://forums.androidcentral.com/android-apps/547447-what-com-facebook-app-manager-why-trying-download-instagram-facebook-messenge.html. [Online; accessed 31-March-2019].

[69] XDA-Developers Forum (Galaxy Note 4). com.facebook.appmanager. https://forum.xda-developers.com/note-4/themes-apps/com-facebook-appmanager-t2919151. [Online; accessed 31-March-2019].

[70] Your Data Is Our Data: A Truecaller Breakdown. https://techcabal.com/2018/05/02/your-data-is-our-data-a-truecaller-breakdown/. [Online; accessed 31-March-2019].

[71] Aafer, Y., Zhang, N., Zhang, Z., Zhang, X., Chen, K., Wang, X., Zhou, X., Du, W., and Grace, M. Hare Hunting In The Wild Android: A Study On The Threat Of Hanging Attribute References. In *Proceedings of the ACM Conference on Computer and Communication Security (CCS)* (2015).

[72] Aafer, Y., Zhang, X., and Du, W. Harvesting Inconsistent Security Configurations In Custom Android ROMs Via Differential Analysis. In *Proceedings of the USENIX Security Symposium* (2016).

[73] Arzt, S., Rasthofer, S., Fritz, C., Bodden, E., Bartel, A., Klein, J., Le Traon, Y., Octeau, D., and McDaniel, P. Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. *Proceedings of the ACM Special Interest Group on Programming Languages (SIGPLAN)* (2014).

[74] Au, K. W. Y., Zhou, Y. F., Huang, Z., and Lie, D. PScout: Analyzing The Android Permission Specification. In *Proceedings of the ACM Conference on Computer and Communication Security (CCS)* (2012).

[75] Dittrich, D., and Kenneally, E. The Menlo Report: Ethical principles guiding information and communication technology research. *US Department of Homeland Security* (2012).

[76] DR WEB. Trojan preinstalled on Android devices infects applications' processes and downloads malicious modules. http://news.drweb.com/news/?i=11390&lng=en. [Online; accessed 31-March-2019].

[77] FELT, A. P., CHIN, E., HANNA, S., SONG, D., AND WAGNER, D. Android Permissions Demystified. In *Proceedings of the ACM Conference on Computer and Communication Security (CCS)* (2011).

[78] FELT, A. P., WANG, H. J., MOSHCHUK, A., HANNA, S., AND CHIN, E. Permission Re-Delegation: Attacks And Defenses. In *Proceedings of the USENIX Security Symposium* (2011).

[79] IKRAM, M., VALLINA-RODRIGUEZ, N., SENEVIRATNE, S., KAAFAR, M. A., AND PAXSON, V. An analysis of the privacy and security risks of android vpn permission-enabled apps. In *Proceedings of the Internet Measurement Conference (IMC)* (2016).

[80] JIANG, Y. Z. X., AND XUXIAN, Z. Detecting Passive Content Leaks And Pollution In Android Applications. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)* (2013).

[81] JOHNSON, RYAN AND STAVROU, ANGELOS AND BENAMEUR, AZZEDINE. All Your SMS & Contacts Belong to ADUPS & Others. https://www.blackhat.com/docs/us-17/wednesday/us-17-Johnson-All-Your-SMS-&-Contacts-Belong-To-Adups-&-Others.pdf. [Online; accessed 31-March-2019].

[82] LI, L., BISSYANDÉ, T. F., KLEIN, J., AND LE TRAON, Y. An investigation into the use of common libraries in android apps. In *Proceedings of the International Conference on Software Analysis, Evolution, and Reengineering (SANER)* (2016).

[83] PAN, E., REN, J., LINDORFER, M., WILSON, C., AND CHOFFNES, D. Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications. *Proceedings of the Privacy Enhancing Technologies Symposium (PETS) 2018*.

[84] RAZAGHPANAH, A., NITHYANAND, R., VALLINA-RODRIGUEZ, N., SUNDARESAN, S., ALLMAN, M., KREIBICH, C., AND GILL, P. Apps, Trackers, Privacy, and Regulators: A Global Study of the Mobile Tracking Ecosystem. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)* (2018).

[85] RAZAGHPANAH, A., VALLINA-RODRIGUEZ, N., SUNDARESAN, S., KREIBICH, C., GILL, P., ALLMAN, M., AND PAXSON, V. Haystack: In situ mobile traffic analysis in user space. *arXiv preprint arXiv:1510.01419* (2015).

[86] REN, J., LINDORFER, M., DUBOIS, D. J., RAO, A., CHOFFNES, D., AND VALLINA-RODRIGUEZ, N. Bug Fixes, Improvements,... and Privacy Leaks.

[87] REYES, I., WIJESEKERA, P., REARDON, J., ON, A. E. B., RAZAGHPANAH, A., VALLINA-RODRIGUEZ, N., AND EGELMAN, S. "Won't Somebody Think of the Children?.Examining COPPA Compliance at Scale. *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)* (2018).

[88] VALLINA-RODRIGUEZ, N., AMANN, J., KREIBICH, C., WEAVER, N., AND PAXSON, V. A Tangled Mass: The Android Root Certificate Stores. In *Proceedings of the International Conference on Emerging Networking EXperiments and Technologies (CoNEXT)* (2014).

[89] VALLINA-RODRIGUEZ, N., SHAH, J., FINAMORE, A., GRUNENBERGER, Y., PAPAGIANNAKI, K., HADDADI, H., AND CROWCROFT, J. Breaking for commercials: characterizing mobile advertising. In *Proceedings of the Internet Measurement Conference (IMC)* (2012).

[90] WANG, H., LIU, Z., LIANG, J., VALLINA-RODRIGUEZ, N., GUO, Y., LI, L., TAPIADOR, J., CAO, J., AND XU, G. Beyond Google Play: A Large-Scale Comparative Study of Chinese Android App Markets. In *Proceedings of the Internet Measurement Conference (IMC)* (2018).

[91] WEI, F., ROY, S., OU, X., AND ROBBY. Amandroid: A Precise and General Inter-component Data Flow Analysis Framework for Security Vetting of Android Apps. In *Proceedings of the ACM Conference on Computer and Communication Security (CCS)* (2014).

[92] WU, D., GAO, D., CHANG, R. K. C., HE, E., CHENG, E. K. T., , AND DENG, R. H. Understanding Open Ports In Android Applications: Discovery, Diagnosis, And Security Assessment. *Proceedings of the Network and Distributed System Security Symposium (NDSS)* (2019).

[93] ZHAUNIAROVICH, Y., AND GADYATSKAYA, O. Small Changes, Big Changes: An Updated View On The Android Permission System. In *Research in Attacks, Intrusions, and Defenses* (2016).

[94] ZHOU, X., LEE, Y., ZHANG, N., NAVEED, M., AND WANG, X. The Peril Of Fragmentation: Security Hazards In Android Device Driver Customizations. In *IEEE Symposium on Security and Privacy (SP)* (2014).

## APÉNDICE

### *A. Distribución de la base de usuarios*

La Tabla IX describe nuestra distribución geográfica de la base de usuarios.

| País (N=130) | Muestras | Vendedores | | Vendedores compartir |
|---|---|---|---|---|
| | | Total | Único | |
| Estados Unidos | 12 % | 36 | 11 | 17 % |
| España | 6 % | 24 | 3 | 11 % |
| Indonesia | 6 % | 26 | 7 | 12 % |
| Italoa | 5 % | 15 | 6 | 7 % |
| Reino Unido | 4 % | 19 | 6 | 9 % |
| Mexico | 3 % | 17 | 3 | 8 % |
| Tailanda | 3 % | 28 | 12 | 13 % |
| Alemania | 3 % | 21 | 2 | 10 % |
| Bélgica | 2 % | 17 | 4 | 8 % |
| Países Bajos | 2 % | 16 | 2 | 8 % |
| **Países total** | **130** | **—** | **214** | |

Cuadro IX: Distribución geográfica de nuestros usuarios. Solo se muestran los 10 principales países.

## *B. Custom permissions*

La Tabla X informa un subconjunto de permisos personalizados definidos por proveedores de dispositivos, MNO, servicios de terceros y conjunto de chips fabricantes.

| PERMISOS FABRICANTES | | | |
|---|---|---|---|
| **Paquete** | **Firma** | **Vendedor(es)** | **Permiso** |
| com.sonyericsson.facebook.proxylogin | Sony Ericsson (SE) | Sony | *com.sonyericsson.permission.FACEBOOK* |
| com.sonymobile.twitter.account | Sony Ericsson (SE) | Sony | *com.sonymobile.permission.TWITTER* |
| android | Sony Ericsson (SE) | Sony | *com.sonymobile.googleanalyticsproxy.permission.GOOGLE_ANALYTICS* |
| com.htc.socialnetwork.facebook | Android (TW) | HTC | **.permission.SYSTEM_USE* |
| com.sonymobile.gmailreaderservice | Sony Ericsson (SE) | Sony | *com.sonymobile.permission.READ_GMAIL* |
| com.sec.android.daemonapp | Samsung Corporation (KR) | Samsung | **.ap.accuweather.ACCUWEATHER_DAEMON_ACCESS_PROVIDER* |
| android | Lenovo (CN) | Lenovo | *android.permission.LENOVO_MDM* |
| com.asus.loguploaderproxy | AsusTek (TW) | Asus | *asus.permission.MOVELOGS* |
| com.miui.core | Xiaomi (CN) | Xiaomi | *miui.permission.DUMP_CACHED_LOG* |
| android | Samsung (KR) | Samsung | *com.sec.enterprise.knox.KNOX_GENERIC_VPN* |
| com.sec.enterprise.permissions | Samsung (KR) | Samsung | *android.permission.sec.MDM_ENTERPRISE_VPN_SOLUTION* |
| com.android.vpndialogs | Meizu (CN) | Meizu | *com.meizu.permission.CONTROL_VPN* |

| PERMISOS MNO | | | |
|---|---|---|---|
| **Paquete** | **Firma** | **Vendedor(es)** | **Permiso** |
| com.android.mms | ZTE | T-Mobile US | *com.tmobile.comm.RECEIVE_METRICS* |
| com.lge.ipservice | LG | T-Mobile US | *com.tmobile.comm.RECEIVE_METRICS* |
| hr.infinum.mojvip | Infinum (HR) [40] | H1 Croatia | *hr.infinum.mojvip.permission.RECEIVE_ADM_MESSAGE* |
| com.locationlabs.cni.att | AT&T (US) | AT&T (US) [47] | *com.locationlabs.cni.att.permission.BROADCAST* |
| com.asurion.android.verizon.vms | Asurion (US) [17] | Verizon (US) | *com.asurion.android.verizon.vms.permission.C2D_MESSAGE* |
| jp.naver.line.android | Naver (JP) | South Korea Telekom | *com.skt.aom.permission.AOM_RECEIVE* |

| PERMISSIONS DE SERVICOS DE TERCEROS | | | |
|---|---|---|---|
| **Paquete** | **Firma** | **Vendedor(es)** | **Permiso** |
| com.facebook.system | Facebook | Facebook | **.ACCESS* |
| com.amazon.kindle | Amazon | Amazon | *com.amazon.identity.auth.device.perm.AUTH_SDK* |
| com.huawei.android.totemweather | Huawei (CN) | Baidu | *android.permission.BAIDU_LOCATION_SERVICE* |
| com.oppo.findmyphone | Oppo (CN) | Baidu | *android.permission.BAIDU_LOCATION_SERVICE* |
| com.dti.sliide | Logia | Digital Turbine | *com.digitalturbine.ignite.ACCESS_LOG* |
| com.dti.att | Logia | Digital Turbine | *com.dti.att.permission.APP_EVENTS* |
| com.ironsource.appcloud.oobe.wiko | ironSource | ironSource | *com.ironsource.aura.permission.C2D_MESSAGE* |
| com.vcast.mediamanager | Verizon (US) | Synchronoss | *com.synchronoss.android.sync.provider.FULL_PERMISSION* |
| com.myvodafone.android | Vodafone (GR) | Exus | *uk.co.exus.permission.C2D_MESSAGE* |
| com.trendmicro.freetmms.gmobi | TrendMicro (TW) | GMobi | *com.trendmicro.androidmup.ACCESS_TMMSMU_REMOTE_SERVICE* |
| com.skype.rover | Skype (GB) | Skype | *com.skype.android.permission.READ_CONTACTS* |
| com.cleanmaster.sdk | Samsung (KR) | CleanMaster | *com.cleanmaster.permission.sdk.clean* |
| com.netflix.partner.activation | Netflix (US) | Netflix | **.permission.CHANNEL_ID* |

| PERMISSIONS CHIPSET | | | |
|---|---|---|---|
| **Paquete** | **Firma** | **Vendedor(es)** | **Permiso** |
| com.qualcomm.location | ZTE (CN) | Qualcomm | *com.qualcomm.permission.IZAT* |
| com.mediatek.mtklogger | TCL (CN) | MediaTek | *com.permission.MTKLOGGER* |
| com.android.bluetooth | Samsung (KR) | Broadcom | *broadcom.permission.BLUETOOTH_MAP* |

Cuadro X: Ejemplos de permisos personalizados. El comodín * representa el nombre del paquete siempre que el prefijo de permiso y el solapamiento del nombre del paquete.