**MÁSTER EN INVESTIGACIÓN EN CIBERSEGURIDAD**

Trabajo de Fin de Máster

Análisis de troyano Chameleon en Android

Autor: Raúl Núñez García

Tutor: Ángel Manuel Guerrero Higueras

Julio 2024

|  |  |  |
| --- | --- | --- |
| **UNIVERSIDAD DE LEÓN**  **Escuela de Ingenierías Industrial, Informática y Aeroespacial**  **MÁSTER EN INVESTIGACIÓN EN CIBERSEGURIDAD**  **Trabajo de Fin de Máster** | | |
| **ALUMNO:** Raúl Núñez García | | |
| **TUTOR:** Ángel Manuel Guerrero Higueras | | |
| **TÍTULO:** Análisis de troyano Chameleon en Android | | |
| **CONVOCATORIA: j**ulio 2024 | | |
| **RESUMEN:**  En este Trabajo de Fin de Máster (TFM) se ha llevado a cabo un análisis detallado de una muestra de malware de tipo troyano que forma parte de la familia conocida como Chameleon, y que ha sido desarrollado infectar dispositivos de la plataforma Android.  Se ha aplicado un análisis estático del código del malware con una fase de deofuscación previa para hacer el código más legible y fácil de comprender. En este análisis se ha estudiado el código del malware para tratar de extraer una lista de posibles tareas que puede llevar a cabo durante su ejecución y los métodos que utiliza para poder tener éxito en la infección y sus pasos posteriores.  Paralelamente se ha llevado a cabo una fase de análisis dinámico, ejecutando el malware en un entorno controlado con dispositivos emulados, recabando información acerca de su comportamiento y realizando comparativas respecto a los datos extraídos en el análisis estático.  Finalmente, se ha documentado toda la información relevante extraída durante los procesos mencionados, tratando de plasmar el funcionamiento en detalle del malware, su peligrosidad y cómo encaja dentro del ecosistema actual de troyanos y malware en Android. | | |
| **Palabras clave: malware, Android, troyano, infostealer, Chameleon** | | |
| **Firma del alumno:** | **VºBº Tutor:** | **VºBº Cotutor:** |

**Abstract**

In this Master Thesis Project, a detailed analysis of a malware sample has been carried out. It revolves around a malware of trojan type that belongs to the family known as Chameleon, and it has been developed for targeting and infecting Android devices.

Static analysis has been applied over the code of the sample, with a previous phase of deobfuscation to make the code more legible and easier to understand. In this static phase the code has been analyzed in detail to try and understand the tasks the malware could use during its execution on an infected device, as well as the methods it employs to try and achieve success in its objectives.

In parallel, a dynamic analysis phase has been carried out, executing the sample in a controlled environment with emulated devices and compiling all the information derived from the workings and behavior the malware exhibits when running, comparing the information extracted with the one obtained during the static analysis.

Finally, all the information collected during the processes mentioned has been documented to try and showcase the detailed workings of the malware, the dangers it poses to users and how it fits within the current Android malware and trojan ecosystem.

**Índice de contenidos**

[Índice de figuras 4](#_Toc169729326)

[Índice de tablas 5](#_Toc169729327)

[Glosario de términos 6](#_Toc169729328)

[Introducción 9](#_Toc169729329)

[Capítulo 1: Estudio del problema 12](#_Toc169729330)

[1.1 EL CONTEXTO DEL PROBLEMA 12](#_Toc169729331)

[1.2 EL ESTADO DEL ARTE 14](#_Toc169729332)

[1.2.5 Conclusión 23](#_Toc169729333)

[1.3 DEFINICIÓN DEL PROBLEMA 23](#_Toc169729334)

[1.4 MALWARE ELEGIDO 24](#_Toc169729335)

[1.4.2 Tecnologías utilizadas 25](#_Toc169729336)

[Capítulo 2: Gestión del proyecto 32](#_Toc169729337)

[2.1 ALCANCE DEL PROYECTO 32](#_Toc169729338)

[2.2 PLAN DE TRABAJO 33](#_Toc169729339)

[2.2.1 Identificación de tareas 33](#_Toc169729340)

[2.2.2 Estimación de tareas 34](#_Toc169729341)

[2.2.3 Planificación de tareas 35](#_Toc169729342)

[2.3 GESTIÓN DE RECURSOS 35](#_Toc169729343)

[2.3.1 Especificación de recursos 35](#_Toc169729344)

[2.3.2 Presupuesto 36](#_Toc169729345)

[2.4 GESTIÓN DE RIESGOS 39](#_Toc169729346)

[2.4.1 Identificación de riesgos 39](#_Toc169729347)

[2.4.2 Análisis de riesgos 40](#_Toc169729348)

[2.5 LEGISLACIÓN Y NORMATIVA 42](#_Toc169729349)

[Capítulo 3: Solución 44](#_Toc169729350)

[3.1 DESCRIPCIÓN DE LA SOLUCIÓN 44](#_Toc169729351)

[3.2 EL PROCESO DE DESARROLLO 45](#_Toc169729352)

[3.2.1 Identificación de muestra 45](#_Toc169729353)

[3.2.2 Análisis estático 46](#_Toc169729354)

[3.3 RESULTADOS DEL DESARROLLO 48](#_Toc169729355)

[Capítulo 4: Evaluación 49](#_Toc169729356)

[4.1 Identificación de la muestra 49](#_Toc169729357)

[4.1.2 Análisis estático 50](#_Toc169729358)

[Capítulo 5: Conclusión 61](#_Toc169729359)

[5.1 APORTACIONES REALIZADAS 62](#_Toc169729360)

[5.2 AMPLIACIONES FUTURAS 62](#_Toc169729361)

[5.3 PROBLEMAS ENCONTRADOS 62](#_Toc169729362)

[5.4 OPINIONES PERSONALES 62](#_Toc169729363)

[Bibliografía 63](#_Toc169729364)

[Anexo A: Control de versiones 67](#_Toc169729365)

[Anexo B: Seguimiento del proyecto 68](#_Toc169729366)

[B.1 FORMA DE SEGUIMIENTO 68](#_Toc169729367)

[B.2 PLANIFICACIÓN INICIAL 68](#_Toc169729368)

[B.3 PLANIFICACIÓN FINAL 68](#_Toc169729369)

**Índice de figuras**

1 4

**Índice de tablas**

**Glosario de términos**

Malware: también conocido como código dañino, se trata de un programa informático que realiza acciones en un sistema sin el conocimiento ni autorización del usuario y con objetivos maliciosos como puede ser el robo de datos o la toma de control del sistema.

Troyano: tipo de malware que en la mayoría de los casos se hace pasar por legítimo, pero realmente permite a los atacantes tomar el control del dispositivo infectado.

Infostealer: tipo de malware diseñado con el objetivo de realizar una exfiltración de datos del dispositivo infectado para obtener información que pueda resultar de interés a los atacantes.

Keylogger: tipo de malware que destaca por su funcionalidad de registro de teclas utilizadas, permitiendo la exfiltración de las cadenas de texto creadas por el usuario del dispositivo del usuario con fines maliciosos como puede ser el fraude. En algunos casos funciona como uno de los componentes de un malware infostealer.

Ofuscación: técnica de modificación del código fuente de un sistema que busca aumentar la dificultad de su comprensión mediante el empleo de estrategias como la reordenación del código o el empleo de cifrado. Este sistema puede ser utilizado por atacantes para tratar de hacer más difícil la labor de análisis estático del malware.

APK: tipo de archivo empaquetado utilizado en Android para las aplicaciones que se distribuyen en el ecosistema. Contiene todos los materiales necesarios para la instalación de una aplicación, incluyendo su código, recursos y certificados. Se construye a partir de código fuente en Java o Kotlin.

ART: acrónimo de Android Runtime, es la máquina virtual que utilizan los dispositivos Android. Es el sistema sucesor de Dalvik y hace las veces de máquina virtual Java, pero cuenta con su propio sistema de bytecode. Realiza la traducción entre el bytecode generado y las instrucciones nativas del dispositivo.

Kernel: núcleo del sistema operativo de un sistema, sirviendo de interfaz entre el hardware y los procesos que se ejecutan. Suele ser el componente del sistema que tiene el máximo privilegio otorgado, y por lo tanto control total sobre las operaciones a realizar.

DEX: iniciales de Dalvik Executable, formato de bytecode que utilizan las máquinas virtuales de ART y Dalvik de las aplicaciones Android. El compilador DEX es el encargado de traducir el código fuente Java o Kotlin en DEX.

JAR: formato contenedor Java Archive, que sirve para empaquetar clases en lenguaje Java, así como sus recursos y metadatos. Se utiliza para la distribución de aplicaciones Java. Puede ser de tipo ejecutable, en cuyo caso indica una clase de entrada en su manifiesto.

APT: acrónimo en inglés de Advanced Persistent Thread. Hace referencia a grupos de atacantes informáticos de gran sofisticación y poseedores de cantidad de recursos para realizar sus operaciones. Suelen ser grupos con financiación y apoyo de grandes organizaciones, algunas de ellas a menudo estados.

IoC: indicadores de compromiso en español, son conjunto de datos y evidencias generados por una actividad realizada que indican las características de una amenaza. Pueden englobar ficheros creados, procesos ejecutados, dominios o direcciones IP contactadas entre otros.

MD5: algoritmo criptográfico de reducción que genera un valor de 128 bits a partir de una entrada de valor arbitrario. Se utiliza a entre otros a modo de checksum para verificar la integridad de los datos.

DDoS: ataque distribuido de denegación de servicio, consistente en la sobrecarga de una infraestructura informática como puede ser un sitio web o un servidor y provocando problemas de disponibilidad en este. En gran parte de los casos, estos ataques utilizan redes de dispositivos infectados y controlados remotamente, conocidos como botnets, para llevar a cabo las ofensivas.

Cryptojacking: tipo de delito informático en el que los dispositivos infectados son utilizados por los atacantes para el minado de criptomonedas sin el conocimiento de la víctima que posee el dispositivo y con el fin de obtener un beneficio económico.

DNS: sistema de nombres de dominio jerárquico y distribuido que establece un sistema de nombres y direcciones para dispositivos que forman parte de Internet. Delega la responsabilidad de gestionar nombres y relaciones para cada dominio a sus servidores DNS autoritativos correspondientes.

API: también conocida como interfaz de programación, es un grupo de definiciones y protocolos que sirven para la integración de diferentes componentes de software.

C&C: del inglés Command & Control, también conocido como C2, alude a la infraestructura utilizada por los atacantes para controlar los equipos infectados de manera remota a través de una comunicación entre los servidores del C2 y los dispositivos.

**Introducción**

**PLANTEAMIENTO DEL PROBLEMA**

**OBJETIVOS**

El objetivo principal de este trabajo es el análisis en profundidad de una muestra de malware de tipo troyano desarrollado para plataformas Android. Los subobjetivos que se esperan cumplir mediante la realización del proyecto son los siguientes:

O1 – Enmarcar el malware elegido y la familia de la que procede dentro del ecosistema actual de troyanos de plataformas Android.

O2 - Realizar un análisis de los mecanismos de distribución del malware.

O3 - Llevar a cabo un análisis estático del código malware para tratar de comprender sus capacidades y modus operandi.

O3 – Realizar un análisis dinámico del malware en un entorno controlado, que permita observar cómo se comporta una vez instalado y ejecutado en un dispositivo infectado

O4 – Comparar los datos obtenidos en el análisis estático con los extraídos en la fase dinámica para confirmar el comportamiento esperado y extraer posibles diferencias en el caso de que estas sean identificadas.

O5 – Obtener una serie de indicadores de compromiso y procedimientos que permitan identificar el malware en el caso de encontrarse este presente en un dispositivo.

**METODOLOGÍA**

El trabajo y la gestión del proyecto se llevarán a cabo siguiendo de manera aproximada la metodología ágil (también conocida como agile), ya que, pese a que existe una clara demarcación entre algunas de las fases debido a su diferente naturaleza, no se trata de unidades aisladas. Los datos obtenidos en cada una de las etapas serán de ayuda en otras fases, permitiendo ampliar de manera continua el conocimiento acerca del funcionamiento del malware. El análisis estático y dinámico se complementarán entre sí, aportando información que permitirá progresar globalmente. No obstante, es importante que no se ha seguido completamente la modalidad con sprints claramente definidos y entregables al final de cada uno como se indica en algunas variantes de agile como puede ser Scrum. Las razones de esto son, por un lado, la dificultad de concretar plazos para ciertas tareas al tratarse de un proyecto que cuenta con un componente importante de investigación, y por lo tanto de varias incógnitas que impiden estimar tiempos con una alta precisión y, por otro lado, debido a la complejidad de compatibilizar el trabajo en el proyecto de manera totalmente definida con el horario laboral del alumno. Sin embargo, se ha tratado en la medida de lo posible de seguir el modelo y su filosofía de adaptación a cambios e iteración frecuente.

Tras la planificación inicial, las etapas del proyecto quedan definidas de la siguiente manera:

1. Diseño de los análisis y planificación de las fases que los compondrán
2. Realización de los análisis y pruebas.
3. Documentación de los resultados obtenidos.

**ESTRUCTURA DEL TRABAJO**

Esta memoria tiene una estructura dividida en capítulos, y cuenta en su parte final con anexos relativos a temas generales del proyecto. A continuación, se muestra un listado de cada uno de ellos, así como una descripción detallada de los contenidos que componen cada ítem:

* Introducción. Planteamiento del problema a resolver, resumen de objetivos, metodología de trabajo y presentación de la estructura de la memoria.
* Capítulo 1. Investigación previa realizada, estado del arte en la temática del proyecto y el proceso de selección del malware a analizar.
* Capítulo 2. Planificación temporal creada inicialmente, junto con presupuestos, la gestión de riesgos y la materia de legislación aplicable al proyecto.
* Capítulo 3. Organización de los análisis a realizar, siendo estos el análisis estático y el dinámico. Presentación del método de trabajo a utilizar durante la realización de las actividades y posibles problemas a encontrar, así como soluciones o mitigaciones previstas.
* Capítulo 4. Presentación de los resultados obtenidos a raíz de los análisis aplicados. Comparativa entre los resultados de las dos fases y corroboración de las coincidencias observadas, además de las diferencias que puedan producirse.
* Capítulo 5. Aportaciones realizadas a través de los resultados obtenidos, al igual que comentario y contextualización de estos. En otro orden de cosas, también se hará una enumeración de los problemas encontrados durante la realización del proyecto, así como posibles futuras mejoras o ampliaciones a aplicar.
* Anexo A. Documentación del progreso realizado y materiales generados, control de versiones del trabajo.
* Anexo B. Seguimiento del proyecto y cambios efectuados respecto a las planificaciones iniciales.

**Capítulo 1: Estudio del problema**

## 1.1 EL CONTEXTO DEL PROBLEMA

En este proyecto se ha llevado a cabo un análisis de una muestra de malware Android de tipo troyano perteneciente a la familia Chameleon. Se han empleado aproximadamente 3 meses en llevar a cabo la tarea incluyendo las investigaciones iniciales, ya que, a pesar de ser un tema cercano a las temáticas del máster, no se cuentan con conocimientos extensos en profundidad para la realización del análisis, especialmente para plataformas móviles como Android. No obstante, algunas asignaturas del título han servido para sentar las bases y permitir avanzar más rápidamente, así como para familiarizarse con algunas de las herramientas a utilizar.

Para llevar a cabo un análisis en profundidad que identifique los vectores de ataque del malware, así como sus procedimientos y objetivos, se han de realizar los siguientes pasos:

* Investigación de la literatura existente respecto al tipo de malware elegido y las familias relacionadas, con el fin de conocer su naturaleza y los procedimientos más habituales que utilizan.
* Análisis estático de la muestra obtenida, que permita entender el funcionamiento interno del malware y las tareas que puede realizar, así como código que pese a no ser necesariamente ejecutado en las pruebas dinámicas también forma parte de la amenaza, pudiendo ser ejecutado bajo ciertas condiciones especiales.
* Análisis dinámico del malware en un entorno controlado, que posibilite visualizar y obtener datos acerca del funcionamiento del troyano en un dispositivo Android, que en este caso será una versión virtualizada del sistema.
* Análisis de los resultados obtenidos en las fases anteriores, determinando las amenazas que presenta el malware para los sistemas Android, por qué tiene éxito en sus ataques y cuáles son los signos o indicadores de su presencia en un dispositivo.

## 1.2 EL ESTADO DEL ARTE

En esta sección se hará una introducción al malware desde la perspectiva de Android resultado de la evaluación y revisión de literatura realizada, dividiéndose en cinco subsecciones según la temática, comenzando por una visión general del malware para la plataforma.

**Malware en Android**

La plataforma Android surgió en 2008 con la llegada al mercado de los primeros dispositivos móviles inteligentes (conocidos popularmente como smartphones). Está basada en el kernel Linux y es un proyecto de código abierto liderado por Google, que cuenta hasta el momento con 21 versiones diferentes y una gran fragmentación en lo relativo a estas [X]. Este es uno de los puntos a destacar desde el punto de vista de la seguridad, ya que esta división con dispositivos que cuentan con versiones significantemente diferentes provoca que existan distintos niveles de protección y mitigaciones en función de la fecha de publicación de la versión concreta de Android, así como de las posteriores actualizaciones de seguridad. Asimismo, existen diferentes variantes de Android dentro de las propias versiones, ya que muchos fabricantes optan por personalizar Android para sus dispositivos con el objetivo de ofrecer características adicionales a los usuarios de sus dispositivos u optimizar el sistema operativo para aprovechar de manera más eficiente los recursos de los que estos disponen. Pese a que el sistema de Google es el más popular, grandes fabricantes como Samsung, Xiaomi o Huawei con su propio sistema Android, lo que aumenta aún más la fragmentación y también supone en algunos casos cambios en las políticas y medidas de seguridad aplicadas.

Hoy en día el sistema Android en sus diferentes versiones engloba a la mayoría de los dispositivos móviles en el mercado, ya que cuenta con una cuota superior al 70%. Este dato, aunado a la fragmentación previamente mencionada, hace de la plataforma Android el objetivo número uno de los atacantes en lo concerniente a sistemas móviles, con iOS en la segunda posición a gran distancia en términos de volumen de ataques.

Pese a la relativa irrupción reciente en el mundo de la informática, Android ha sido un blanco de malware desde sus inicios. En 2010 surgió el primer malware específico para Android, aunque ya existía malware para sistemas operativos móviles desde 2004 [X] y desde entonces el número de amenazas en Android ha crecido rápidamente en comparación con otras plataformas, aumentando principalmente en volumen, pero también en sofisticación. Los fabricantes tratan de incluir nuevas medidas de prevención y reducción de riesgos, así como realizar campañas de prevención, algunas de ellas en colaboración con organismos públicos.

No obstante, el malware en Android sigue siendo una actividad altamente lucrativa, no solamente en términos económicos, si no también geopolíticos, como ha sido demostrado por algunos de los últimos ataques de APTs, en los que Android ha sido la plataforma y la finalidad es la exfiltración de datos de carácter confidencial o sensible.

**Troyanos en Android**

El malware de tipo troyano es uno de los más populares en términos globales sin tener en cuenta la plataforma, y el sistema Android sigue el patrón. Se estima que entorno a un 30% del malware de Android se puede encuadrar dentro de esta tipología.

Dentro de las muestras categorizadas como troyanos se pueden encontrar dos modalidades cuando hablamos de Android. La primera de ellas se trata de una aplicación que simula ser legítima, y siendo en esencia una versión modificada de la aplicación a la que trata de suplantar. En esta versión alterada se encuentra el código malicioso que permite al troyano la realización de sus acciones maliciosas. La tasa de éxito de esta tipología de troyano viene dada por su capacidad para emular las características de la aplicación suplantada, tratando de ser una copia casi exacta y en muchos casos contando con gran parte del código de la aplicación original. Para tratar de aumentar las posibilidades de infección, los atacantes suelen elegir aplicaciones populares y conocidas para suplantar.

Por otro lado, existen troyanos dentro de aplicaciones que simulan ser benignas sin suplantar a otras existentes, como pueden ser utilidades de sistema o plataformas de juegos que cuentan con funcionalidades legítimas en algunos casos, pero que siempre contienen componentes maliciosos.

Al basarse Android en un sistema aplicaciones y siendo gran parte de ellas ajenas al sistema original que se instala en el dispositivo, los troyanos se adaptan perfectamente a este modelo y son una de las tipologías de malware más eficaces y con mayor capacidad lucrativa para los grupos de atacantes.

**Métodos de distribución**

Uno de los factores importantes a tener en cuenta en los troyanos en Android es el canal de distribución. Los atacantes tratan de minimizar el nivel de sospecha por parte de los usuarios para facilitar la instalación de las aplicaciones maliciosas, y para ello utilizan diferentes mecanismos.

El primero de ellos y uno que aumenta la credibilidad del troyano es su distribución directa mediante canales oficiales de aplicaciones. El principal de estos canales es la Play Store de Google, que cuenta en la actualidad con más de 3 millones de aplicaciones para el sistema Android. En el sistema Android por defecto, esta es la única tienda de aplicaciones disponible, aunque algunos fabricantes utilizan también plataformas propias para la distribución de aplicaciones, como es el caso de Samsung con su Galaxy Store o la Appstore de los dispositivos Amazon.

Además, existen canales de distribución alternativos no asociados a los fabricantes, pero que también cuentan con un gran volumen de usuarios debido a las características particulares que ofrecen, como por ejemplo F-Droid [X] , que distribuye únicamente aplicaciones de código abierto o Aptoide [X] , un mercado secundario popular debido a que almacena y permite la descarga de aplicaciones que originalmente son de pago de manera gratuita, y también facilita la descarga de versiones específicas de aplicaciones que ya no están disponibles a través de los canales oficiales debido a su antigüedad.

Asimismo, una de las opciones del sistema Android es la instalación de aplicaciones directamente por parte de los usuarios si estos cuentan con el archivo APK y lo transfieren al dispositivo. Pese a que por defecto la seguridad del sistema previene la instalación de aplicaciones que provienen de fuentes desconocidas, esta prevención puede ser desactivada por los usuarios de manera sencilla en los ajustes del sistema. Esto supone un vector adicional de entrada de aplicaciones al sistema, y en este caso se cuenta con un nivel de protecciones bajo relativo a las otras opciones de distribución mencionadas previamente.

Por defecto, la tienda Play Store de Google y los mercados de los fabricantes mayoritarios cuentan con comprobaciones y análisis que son efectuados en las aplicaciones previa a su publicación en el catálogo del mercado. Esto minimiza la cantidad de aplicaciones maliciosas que logran la publicación, pero sigue siendo una cantidad importante de ellas la que logra pasar las barreras y establecerse en los mercados oficiales, con millones de descargas por parte de usuarios [X].

Además, Google cuenta con funcionalidades de seguridad adicionales en el caso de que la aplicación ya se encuentre instalada en el sistema. Con el objetivo de identificar PHAs (término de Google para englobar aplicaciones que pueden provocar daños al dispositivo o sus datos, que proviene del inglés Potentially Harmful Applications [X]) Google ha desarrollado el sistema Google Play Protect, que realiza escaneos diarios de todas las aplicaciones presentes en el dispositivo, además de intentar prevenir instalaciones de aplicaciones externas a Google Play si estas son detectadas como maliciosas en función de las muestras dañinas de las que poseen información previa. Como medidas adicionales, este servicio Play Protect permite a los usuarios la petición de escaneos suplementarios, y también tiene un modo de funcionamiento que permite los escaneos cuando el dispositivo carece de conexión a Internet.

En definitiva, Google (y en menor medida otros de los principales fabricantes) cuentan con numerosos sistemas de prevención y defensa para tratar de combatir a las aplicaciones maliciosas. No obstante, esto no impide que sigan apareciendo casos en los que los atacantes son capaces de sobrepasarlas y cumplir sus objetivos maliciosos, bien a través de canales oficiales, bien mediante el tráfico de APKs en mercados secundarios. Además, las medidas de seguridad no cubren todos los casos, ya que algunas funcionalidades, como es el caso de los permisos de accesibilidad que se verán en detalle a continuación, quedan fuera de las pruebas y comprobaciones de seguridad realizadas.

**Sistema de permisos de accesibilidad**

El abuso de los permisos es uno de los vectores de ataque más comunes por parte de aplicaciones Android, contando muchas veces con la colaboración del usuario, que desconoce qué permisos está realmente otorgando o no presta demasiada atención al proceso inicial al instalar una aplicación. Debido a esto, Android cuenta desde su versión 12 con medidas adicionales de notificación para tratar de alertar a los usuarios en casos como la activación de la cámara o el micrófono del dispositivo, conocidas como “Privacy indicators” [X]. Por esto, algunos de los desarrolladores de aplicaciones maliciosas han modificado sus mecanismos de acceso a estos recursos, haciendo uso de otro tipo de permisos que otorga el sistema con la esperanza de que no alerten al usuario de manera tan repentina y este acepte los permisos sin levantar sospechas.

Algunos de estos permisos son relativos a la accesibilidad, la herramienta de los sistemas operativos móviles (y sistemas informáticos en general) con la finalidad de facilitar y hacer lo más práctica posible la experiencia de uso para las personas con discapacidades visuales, auditivas, cognitivas o físicas. En el caso de Android, las primeras opciones de accesibilidad fueron introducidas en su versión 1.6 (conocida popularmente como Donut) y se centraron en el lector de pantalla TalkBack, con nuevas características surgiendo en versiones posteriores del sistema, como la navegación por gestos, el control del dispositivo por voz mediante el sistema Voice Access, la generación automática de subtítulos o la conexión con audífonos o implantes cocleares, además de contar con Android Accesibility Suite [X], que reúne la gestión de varios de estos servicios en una única aplicación oficial de Google.

Otro de los mecanismos que los atacantes utilizan en algunos casos es el sistema de administración de dispositivo. Este sistema es una API que Android proporciona desde su versión 2.2 (conocida como Froyo) [X] y que permite a los desarrolladores de aplicaciones manejar a través de estas la gestión de ciertos servicios del dispositivo que por lo general se encuentran fuera del sistema habitual de permisos. Esta funcionalidad está diseñada principalmente para entornos de empresa en los que se busca facilitar la gestión de numerosos dispositivos empresariales de manera controlada y segura, permitiendo la instalación de aplicaciones y el control sobre las políticas de bloqueo de pantalla entre otros. La lista de permisos que otorga este rol de administrador a través de la API varía en función de los requeridos por la aplicación que cuenta con el rol. Por lo tanto, pese a que esta característica puede resultar útil en ciertos casos, también puede ser abusada para tratar de obtener mayores privilegios bajo el pretexto del control seguro del dispositivo. No obstante, para que una aplicación sea asignada como administradora del dispositivo, el usuario ha de aceptar los términos, por lo tanto, nuevamente es la falta de atención o el desconocimiento de los usuarios que abre la puerta a la entrada del malware.

**Análisis de malware en Android**

Pese a que en sus inicios el malware para sistemas móviles era rudimentario en comparación con el existente en los sistemas informáticos más maduros como los ordenadores, con el tiempo los primeros han ido aumentando su nivel de sofisticación.

En muchos de los casos, el análisis de malware para los sistemas móviles, y particularmente Android al ser la plataforma más popular, está basado en métodos y técnicas similares provenientes del análisis de malware tradicional, al ser estrategias que han demostrado ser útiles en la identificación y prevención de software dañino. En ambas plataformas, las principales técnicas son el análisis estático y el análisis dinámico. En primer lugar, hablaremos del análisis estático aplicado a Android.

Pese a utilizar técnicas similares al análisis tradicional de malware, para el análisis estático en Android las herramientas comunes para plataformas como Windows o Linux, como pueden ser los analizadores de cabeceras o ficheros, no son normalmente de gran utilidad. Además, el código fuente del malware rara vez se encuentra disponible, teniendo que ser analizado en su lugar el bytecode que se encuentra dentro de los ficheros APK de las aplicaciones maliciosas, y que en algunos casos también contará con medidas de protección similares a las encontradas en sistemas como Windows, como pueden ser el packing [X] de aplicaciones para la ocultación de sus clases reales y su posterior carga dinámica [X] para evadir la detección estática de manera sencilla, o la ofuscación del propio código para hacerlo más difícil de comprender en el futuro por parte de posibles analistas [X]. Para lidiar con estos casos, se cuenta con una variedad de herramientas específicas para para Android, como pueden ser simplify [X] para tratar de revertir algunos de los métodos de ofuscación más comunes y populares, como pueden ser los aplicados por el programa ProGuard [X], o el unpacker HoseDex2Jar [X]. Una vez obtenido el fichero DEX definitivo que declara las clases que utilizará la aplicación, se puede tratar de obtener código Java a partir de él, haciendo uso de herramientas populares como JADX [X].

No obstante, en el caso de ser superada la fase de medidas de anti-análisis comentadas, y estando en posesión de bien código Java o tratarse de un caso de malware que utiliza librerías nativas compiladas a partir de código C o C++, algunas herramientas que forman parte del arsenal de análisis estático estándar son de gran utilidad, como puede ser Ghidra [X], además de técnicas establecidas en el análisis tradicional como el análisis de dependencias mediante grafos.

En lo relativo al análisis dinámico, Android cuenta con menos opciones para la virtualización de dispositivos que las disponibles para plataformas más establecidas en el mundo del malware como son Windows o Linux. Se distinguen dos tipos de análisis en este caso, conocidos como in-the-box y out-the-box. El primero de ellos hace referencia al análisis en el que se ejecuta la muestra maliciosa y el análisis junto con la recogida de información se produce en el mismo nivel y en el mismo dispositivo. Para llevar a cabo este método se necesita de utilidades que trabajen en el mismo nivel que la máquina virtual Dalvik y tratan de instrumentar procesos internos de Android para obtener información acerca de las modificaciones realizadas por el malware, como por ejemplo es el caso de la herramienta DIVILAR [X]. Un punto en contra de este tipo de métodos que trabajan en un nivel de abstracción igual que el propio malware, es que este último puede realizar comprobaciones para tratar de detectar si el entorno está siendo alterado con fines analíticos y modificar su comportamiento para tratar de pasar desapercibido, o simplemente dejar de funcionar [X].

El segundo método, out-the-box, también conocido como análisis de tipo VM, se trata del uso de emuladores y entornos virtuales aislados (también denominados sandbox) como pueden ser el propio emulador de la suite Android Studio de Google o herramientas de terceros como GenyMotion [X] o Android-x86 [X], realizando el análisis desde un entorno externo en la misma máquina. Pese a que las diferencias entre la emulación y un entorno real son casi imperceptibles, incluyendo la simulación de elementos como el sensor GPS, la cámara, o el lector de huellas dactilares entre otros, existen diferencias y nuevamente el malware puede detectar que está siendo ejecutado en un entorno modificado, respondiendo adecuadamente para tratar de evitar el análisis. Pese a que este segundo método no trabaja en un nivel tan cercano al propio malware, permite la utilización de multitud de herramientas para tratar de instrumentar la ejecución de este, como pueden ser FRIDA [X], herramientas derivadas como objection [X] o medusa [X], además de utilidades como Wireshark para la captura de tráfico o la suite de Burp para realizar operaciones de proxy como el Certificate Pinning [X].

Es importante mencionar que las herramientas de análisis dinámico enumeradas hasta el momento son todas ellas aplicables a entornos locales, y se ha de tener en cuenta que hay también un ecosistema de ofertas en la nube como por ejemplo JoeSandbox [X], que permiten la asignación de espacios aislados en los que ejecutar aplicaciones Android potencialmente maliciosas y recabar información acerca de su comportamiento, que pese a no ofrecer el control ni la granularidad de análisis que se pueden obtener con métodos locale, pueden ser una solución adecuada para un análisis de menor profundidad pero que proporcione datos importantes sobre el malware. No obstante, en este espacio de soluciones online el ecosistema Android carece del número de plataformas dedicadas a otros sistemas como Windows o Linux.

Por último, además de los dos métodos principales de análisis, también se puede considerar un tercero conocido como análisis híbrido, que aúna el análisis estático y el dinámico, ejecutándolos de manera cercana al paralelismo para tratar de explorar todos los caminos de ejecución que puede ofrecer una aplicación y tratar de maximizar el área del código del malware cubierta por el análisis. Herramientas como el módulo A5 [X] o el algoritmo de tipo concólico (que aúna análisis concreto y simbólico) ALLSEQS [X] son algunos ejemplos de este tipo de técnicas.

### 1.2.5 Conclusión

Como se ha expuesto en la sección anterior, el estado actual tanto del malware en Android, específicamente de las aplicaciones de tipo troyano, como también de las herramientas y técnicas que existen para su análisis son ampliamente variadas.

Pese a que el entorno móvil no cuenta todavía con el mismo abanico de posibilidades y recursos que los entornos mayoritarios de escritorio como Windows o Linux, tanto la sofisticación del malware como los procedimientos para su detección se encuentran en el centro de una carrera constante de innovación entre atacantes y analistas del lado de la defensa que hace que el análisis de malware en Android sea un tema de actualidad y gran importancia dentro del mundo de la ciberseguridad.

## 1.3 DEFINICIÓN DEL PROBLEMA

Pese a las medidas de seguridad existentes en Android, una parte importante del malware sigue teniendo éxito en su cometido, con la combinación de distintos tipos de ataques y mecanismos para tratar de obtener privilegios adicionales. En muchos casos el desconocimiento de los propios usuarios y la falta de comprobaciones adicionales provoca las infecciones y los consiguientes comportamientos maliciosos por parte de las aplicaciones. El abuso de los permisos es uno de los vectores que puede utilizar el malware en Android, particularmente aquellos de tipo troyano.

Es importante recabar información para conocer su funcionamiento en detalle, desde la llegada de la aplicación al dispositivo previa a su instalación hasta sus técnicas para conseguir tomar el control de los dispositivos y las acciones que realiza. De toda esta información que se ha de recabar se puede extraer un perfil que identifique al malware bajo análisis, y que permita crear una serie de IoCs que permitan su reconocimiento preventivo de manera automática, la obtención de información acerca de los grupos de atacantes que pueden estar tras su creación y explotación en el caso de que esto sea posible, además de poder utilizar los datos sobre el proceso del malware para petición de permisos y los pasos iniciales que efectúa para así utilizarlos en la realización de campañas de advertencia con finalidad preventiva destinadas a los usuarios.

## 1.4 MALWARE ELEGIDO

Se ha elegido una muestra de la familia de troyanos conocida como Chameleon. Pese a tener el mismo nombre que una botnet surgida en 2013 [X], la familia en la que se centrará el proyecto tiene únicamente como objetivo la infección de dispositivos Android. Chameleon apareció online por primera vez a principios de 2023 [X], y desde entonces ha pasado por diferentes iteraciones, introduciendo nuevos mecanismos para obtener sus objetivos y evadir la detección por los usuarios. Las razones detrás de esta elección son principalmente dos. En primer lugar, se trata de un malware reciente y que se encuentra actualmente en el foco de la prensa generalista debido a la gran cantidad de amenazas que presenta pese a ser un único malware [X]. Asimismo, esta nueva versión es diferente a las anteriores, y también a otros troyanos para sistemas Android, en la manera en que gestiona los parámetros de accesibilidad para hacerse con el control del dispositivo y evitar la protección biométrica para su desbloqueo [X]. Además, pese a ser un malware relativamente popular en estos últimos años, y aunque como se ha mencionado, sí que existen multitud de artículos de carácter general en medios de comunicación reconocidos haciendo referencia a la amenaza que presenta para los usuarios, no se han publicado artículos especializados ni análisis detallados del malware, siendo el único análisis encontrado el desarrollado por la empresa estadounidense de ciberseguridad Cyble en abril de 2023 para una de las nuevas versiones del troyano [X]. Por todas estas razones expuestas se ha decidido finalmente seleccionar esta familia de malware como temática del proyecto, y más concretamente la versión de Chameleon que se hace pasar por la aplicación Google Chrome y que posee el valor hash SHA-256 de 0a6ffd4163cd96d7d262be5ae7fa5cfc3affbea822d122c0803379d78431e5f6. La muestra ha sido obtenida del popular repositorio de malware y publicaciones de ciberseguridad Vx Underground [X].

### 1.4.2 Tecnologías utilizadas

A continuación, se incluyen las tecnologías utilizadas a lo largo de la realización del proyecto, formando parte de ambas fases de análisis del malware. En primer lugar, se comentan las herramientas utilizadas para realizar el análisis estático del código.



* JADX: se trata de una de las herramientas más populares para descompilar código DEX proveniente de un archivo APK Android a código Java, que es significantemente más fácil de comprender. Además, cuenta con una versión con interfaz gráfica de usuario que permite la visualización de proyectos completos, así como el renombrado de clases, métodos y código junto con el añadido de comentarios, siendo de especial utilidad en la reestructuración y anotación de código ofuscado.



* Apktool: herramienta versátil que sirve para desempaquetar archivos APK y obtener fácilmente a través de la línea de comandos todos los ficheros que componen la aplicación, además de permitir realizar tareas adicionales como el reempaquetado una vez se han aplicado modificaciones en el código.



* Apkid: utilidad para la identificación de packers, protectores y ofuscadores en el proceso de creación de una aplicación. Si algún método popular se ha utilizado para la realización de estas acciones de evasión, Apkid informa de ello, además de proporcionar detalles respecto a la versión particular de la herramienta utilizada para la compilación del proyecto. En resumen, Apkid podría considerarse una herramienta similar en cuanto a funcionalidad a PEiD en Windows, pero para la plataforma Android.



* Sublime Text: editor de texto ligero que soporta el procesamiento y la visualización de una gran variedad de formatos. Utilizado principalmente en el proyecto para la toma de notas de carácter misceláneo con el fin de guiar los procesos de análisis y permitir registrar de manera rápida eventos o detalles que podrían ser de utilidad en el futuro.



* Android Studio: suite de desarrollo de aplicaciones Android proporcionada por Google. Además de servir para crear aplicaciones, también sirve para la depuración y el análisis de archivos APK. Cuenta con un sistema de emulación de dispositivos Android que también hace que esta herramienta sea utilizada intensivamente en la fase de análisis dinámico. Al ser una herramienta de Google, está integrada totalmente en el ecosistema Android, y por lo tanto ofrece un gran abanico de opciones a la hora de emular dispositivos Android, permitiendo personalizar parámetros como ubicación, simulación de envío de mensajes SMS o simulación de ubicaciones GPS.



* IntelliJ IDEA: entorno de desarrollo integrado orientado al trabajo con código Java y Kotlin. Es uno de los IDEs más populares debido a su gran número de opciones y rendimiento. En este proyecto se utiliza para la creación de programas auxiliares en Java que sirvan para tareas variadas del análisis, como la desencriptación de valores o la instrumentación de código.

En lo relativo a la sección de análisis dinámico del troyano, las herramientas utilizadas para su realización fueron las siguientes:

* Android Studio: como se ha mencionado en la sección anterior, el emulador de Android Studio es una pieza muy importante en el análisis dinámico, ya que pese a que contiene algunas limitaciones respecto a otros como la falta de acceso superusuario.



* Adb: parte de las herramientas de depuración de Android, permite la conexión de dispositivos Android (físicos o emulados) a los entornos de depuración elegidos. En este caso será utilizado para la conexión a los dispositivos emulados, con la finalidad de descargar y enviar archivos a estos, así como poder conectar consolas para realizar algunas fases del análisis como el Certificate Pinning.



* Genymotion: plataforma de emulación de dispositivos Android utilizada paralelamente a Android Studio para el análisis dinámico. En su versión gratuita las opciones son limitadas, pero permite la creación de dispositivos con permisos de superusuario, lo cual es importante para algunas de las operaciones.



* FRIDA: Sistema de instrumentación dinámica que permite entre otras funciones el análisis y depuración de código Android en tiempo real a través de la inyección de código JavaScript. Esta funcionalidad facilita la extracción de información mientras el malware se está ejecutando, así como la instrumentación para forzar caminos de ejecución determinados.
* Burp Suite: aplicación muy utilizada en el entorno de la ciberseguridad para realizar análisis y tests de penetración. En este proyecto se utiliza su función de proxy (opcionalmente para el proceso de Certificate Pinning en el caso de realizarse), que permite el análisis de las peticiones enviadas entre el servidor de Command&Control y el malware del dispositivo emulado.



* Wireshark: herramienta de análisis de tráfico de red más popular. Se trata de un programa de código abierto que permite la captura de paquetes de red y su posterior disección y análisis. En este caso servirá para tener una visión global del tráfico entre el dispositivo infectado y el servidor de comando y control remoto.



* DB Browser: programa de código abierto con la función de navegador de bases de datos de SQLite, que se utiliza para observar el esquema de la base de datos creada por el malware, así como la información que se almacena en ella.

Por último, se presentan herramientas adicionales utilizadas durante el proyecto sin pertenecer a ninguna de las fases en particular, si no que han sido empleadas para tareas diversas durante todas las fases:



* Homebrew: gestor de paquetes utilizado en el sistema Mac para la instalación de programas y utilidades desde la línea de comandos, facilitando la gestión de dependencias.



* Pip: gestor de paquetes y dependencias Python que simplifica la instalación de utilidades como Apkid o FRIDA.



* HexFiend: visualizador hexadecimal de archivos cuya utilización en el proyecto ha estado ligada a intentos de identificación de tipo de ficheros basados en su firma [X].



* Microsoft Word: procesador de texto utilizado en realización de la memoria.



* Microsoft Excel: procesador de hojas de cálculo utilizado para la generación de los diagramas Gantt de la planificación de tareas.

# Capítulo 2: Gestión del proyecto

En este apartado se especifica cómo se ha realizado la gestión del proyecto, incluyendo la planificación temporal del mismo, los presupuestos de su realización, la gestión de riesgos existentes, así como la legislación aplicable correspondiente.

## 2.1 ALCANCE DEL PROYECTO

El proyecto trata de obtener la mayor cantidad de información posible de la muestra de troyano seleccionada a través de las tareas de análisis que se enumerarán a continuación. Esta información se podría posteriormente aplicar para la prevención de infecciones y la ayuda a la defensa de malware, en concreto el perteneciente a la familia Chameleon.

## 2.2 PLAN DE TRABAJO

Para llevar a cabo la planificación de este proyecto se han estimado una serie de tareas a realizar y una duración aproximada del tiempo que se podría tardar en ejecutarlas. Más adelante, en el [Anexo B](#_B.1__), se mostrará la duración real que han tenido finalmente dichas tareas.

### 2.2.1 Identificación de tareas

Las tareas principales que se han definido para conseguir llegar a la solución propuesta son las siguientes:

* Tarea 1: realizar un estudio de la literatura existente de malwares troyanos en Android para tener una perspectiva general del estado actual de amenazas en la plataforma.
* Tarea 2: elegir una familia concreta de malware y obtener muestras para su posterior análisis, además de estudiar sus medios de distribución habituales y recabar información acerca de posibles métodos de evasión que pueda presentar, con el objetivo de una mejor estimación del trabajo a realizar en las tareas posteriores.
* Tarea 3: realizar un análisis estático de las muestras del malware, con la finalidad de entender el código que lo compone y las tareas que realiza en la medida de lo posible. Sin requerir el descifrado completo de su código en el caso de estar este ofuscado, será importante comprender una cantidad suficiente como para discernir el funcionamiento general del malware y las acciones (maliciosas o no) que realiza en el dispositivo.
* Tarea 4: realizar un análisis dinámico de las muestras, recabando los IoCs que puedan registrarse y observando el comportamiento del malware, comprobando paralelamente que los resultados obtenidos en el análisis estático se corresponden con los patrones observados en el dispositivo infectado.
* Tarea 5: realizar una comparativa entre los resultados obtenidos en ambos análisis, incluyendo toda la información que pueda resultar relevante para los objetivos del proyecto, y también remarcando las diferencias observadas en el caso de haberlas.
* Tarea 6: elaboración del informe técnico que contenga la información recogida durante las tareas previas, maquetación de los datos y redacción de estos en versión final de la memoria.

### 2.2.2 Estimación de tareas

Para la elaboración de este proyecto se estima una duración total de unas 250 horas repartidas a lo largo de 3 meses, lo que equivaldría a unas 2,7 horas de trabajo diario. Es importante remarcar que se trata de un proyecto realizado paralelamente al horario laboral del alumno, lo cual podría afectar significativamente el tiempo empleado en cada tarea. La estimación de la duración de cada una de tareas presentadas se encuentra a continuación:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tarea** | **Días** | **Horas** | **Fecha de inicio** | **Fecha de finalización** |
| Tarea 1 | 15 | 40,5 | 13/3/2024 | 27/3/2024 |
| Tarea 2 | 10 | 27 | 27/3/2024 | 6/4/2024 |
| Tarea 3 | 35 | 54 | 6/4/2024 | 10/5/2024 |
| Tarea 4 | 35 | 40,5 | 6/4/2024 | 10/5/2024 |
| Tarea 5 | 15 | 40,5 | 10/5/2024 | 25/5/2024 |
| Tarea 6 | 15 | 40,5 | 25/5/2024 | 8/6/2024 |

### 2.2.3 Planificación de tareas

Tras la identificación y estimación de las tareas, se ha generado un diagrama Gantt como perspectiva general de las tareas y su duración a lo largo del proyecto completo.



## 2.3 GESTIÓN DE RECURSOS

En este apartado se presentarán en detalle los recursos que son necesarios para la realización del proyecto junto con sus características principales.

### 2.3.1 Especificación de recursos

* **Recursos materiales**: como materiales para el proyecto, se necesitará de un ordenador portátil para realizar los análisis y trabajar con los resultados. El dispositivo utilizado es un MacBook Pro de 2023 que cuenta con las siguientes características:
  + CPU Apple M2 Pro ARM
  + Pantalla de 14 pulgadas resolución Retina
  + 16GB de memoria RAM
  + 512GB de almacenamiento SSD
  + GPU integrada M2 de 10 núcleos

En lo relativo al software utilizado, ninguna herramienta ha supuesto un coste de licencia, al tratarse en todos los casos de programas gratuitos o, en el caso de ser programas de pago, se ha utilizado la licencia básica, educativa o de demostración que carecen de costes asociados. Por otro lado, se incurrirá en gastos de luz y material de oficina para realizar las tareas necesarias.

* **Recursos humanos:** se contará únicamente con un analista de malware para llevar a cabo las operaciones necesarias. Esta persona estará encargada de realizar los análisis aplicados a la muestra seleccionada, recabar la información que considere relevante y posteriormente organizarla y plasmarla en el informe.

### 2.3.2 Presupuesto

A continuación, se desglosará el presupuesto del proyecto, incluyendo tanto los recursos materiales utilizados como los recursos humanos. En lo referente a gastos de espacios físicos, se han considerado los gastos de luz y material de oficina básico, sin incluir los referentes a un espacio físico concreto.

**Coste de recursos materiales**

* **Ordenador portátil:** su precio de compra fueron 1.850€. Se ha de tener en cuenta que pese a ser utilizado en el proyecto durante únicamente 3 meses, se prevé amortizar el dispositivo durante 4 años. Por lo tanto, su coste final en lo relativo al proyecto asciende a 115,62€.
* **Gastos de luz:** con una utilización de 2,7 horas al día del dispositivo, se supone que el consumo medio del ordenador elegido es de 70 vatios y un precio de 0,05 €/kWh, con lo que obtenemos un precio final para los 3 meses de 12,15€.
* **Material de oficina:** gastos de material diverso como folios, bolígrafos y demás papelería suponen un gasto final de 15€.

**Coste de recursos humanos**

* **Analista de malware:** dado que el analista es un ingeniero informático de formación, se aplicarán las Bases y tipos de Cotización de la Seguridad Social para 2024 para el grupo de “Ingenieros y Licenciados” [X]. Durante este proceso es importante distinguir que se trata de un contrato a tiempo parcial. En primer lugar, los tipos de cotización aplicables serán los siguientes:

|  |  |
| --- | --- |
| **Tipo de cotización** | **Porcentaje (%)** |
| Contingencias comunes del trabajador | 4,70 |
| Cuota de desempleo del trabajador | 1,60 |
| Cuota de formación del trabajador | 0,10 |
| Contingencias comunes de la empresa | 23,60 |
| Cuota de desempleo de la empresa | 6,70 |
| Cuota de formación de la empresa | 0,60 |
| Fondo de Garantía Salarial (FOGASA) | 0,20 |
| Accidentes de trabajo y enfermedades profesionales | 1,55 |
| **Total** | **39,05** |

Teniendo en cuenta que se trata de un contrato de trabajo a tiempo parcial, la base mínima horaria es de 11,13€. Podemos suponer para el caso del analista de nuestro proyecto que su salario ascenderá a 18€ por hora trabajada. Aplicando las cuotas mencionadas previamente obtenemos los siguientes datos:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Rol** | **Horas** | **Coste (€/hora)** | **Coste con cuotas incluidas (€/hora)** | **Total (€)** |
| Analista de malware | 250 | 18 | 25,029 | 6.257,25 |

**Beneficio industrial**

Para obtener el beneficio del proyecto, teniendo en cuenta la Ley 9/2017 de 8 de noviembre del BOE, este será un 6% del coste previo a la aplicación del IVA. Por lo tanto, el beneficio será el siguiente:

|  |  |
| --- | --- |
| **Tipo de recursos** | **Coste (**€) |
| Recursos materiales | 142,77 |
| Recursos humanos | 6.257,25 |
| Subtotal | 6.400,02 |
| Beneficio industrial | 384,0012 |

**Coste total**

Finalmente, tras haber calculado los costes y el beneficio industrial obtenido, se aplica el tipo general del IVA (21%) para obtener el coste total del proyecto. La tabla de presupuestos en su versión final es la siguiente:

|  |  |
| --- | --- |
| **Concepto** | **Coste (**€) |
| Recursos materiales | 142,77 |
| Recursos humanos | 6.257,25 |
| Beneficio industrial | 384,0012 |
| Subtotal | 6.784,0212 |
| IVA | 1424,64 |
| **Coste total** | **8.208,66** |

## 2.4 GESTIÓN DE RIESGOS

Un análisis previo de los riesgos ha de ser realizado previamente a la realización del proyecto. En este apartado se incluirán los riesgos identificados, incluyendo medidas de mitigación en el caso de que estos riesgos se manifiesten.

### 2.4.1 Identificación de riesgos

De acuerdo con la Guía de los fundamentos para la dirección de proyectos, conocida como PMBOK [X], los riesgos que se pueden presentar en el proyecto se dividen en cuatro categorías diferentes en función de su naturaleza. Teniendo en cuenta esta división en categorías, se han identificado los siguientes riesgos asociados con cada una de ellas:

* Riesgos técnicos: son aquellos que hacen referencia a las tecnologías de las que hace uso el proyecto, la calidad de este o la complejidad de las tareas a realizar.
  + Riesgos aplicables al proyecto: problemas de compatibilidad con las tecnologías a utilizar, complicaciones para la obtención de muestras de malware que cumplan las características deseadas, dificultades a la hora de realizar los análisis planificados debido a problemas con las tareas necesarias (código encriptado, ofuscación difícil de revertir, etc.).
* Riesgos externos: relacionados con agentes externos al proyecto.
  + Riesgos aplicables al proyecto: problemas de acceso a Internet.
* Riesgos de la organización: aquellos que engloban a los recursos del proyecto o su financiación.
  + Riesgos aplicables al proyecto: fallo de hardware del equipo que se utilizará para el análisis que provoquen su indisponibilidad para realizar las tareas.
* Riesgos de la dirección de proyectos: referentes a la gestión del proyecto desde el punto de vista de las tareas a realizar, sus plazos de duración y las dependencias entre estos.
  + Riesgos aplicables al proyecto: estimaciones erróneas de la duración de las tareas y fallos de planificación que provoquen alteraciones en las fechas propuestas.

### 2.4.2 Análisis de riesgos

Dados los riesgos identificados, se procede a proponer una serie de medidas para tratar de solventar los problemas que causen en el caso de que ocurran, presentadas en la tabla a continuación:

|  |  |  |
| --- | --- | --- |
| **Categoría** | **Riesgo** | **Medidas de subsanación** |
| Riesgos técnicos | Problemas de compatibilidad de tecnologías | Elección de tecnologías alternativas conocidas que pese a no ser idénticas obtengan resultados similares |
| Riesgos técnicos | Complicaciones para la obtención de muestras de malware que cumplan las características propuestas | Búsqueda en repositorios alternativos de malware |
| Riesgos técnicos | Dificultades a la hora de realizar los análisis (ofuscación no reversible, datos encriptados, etc.). | Realización en paralelo de otros tipos de análisis para tratar de obtener más información sobre el malware que permita una solución alternativa |
| Riesgos externos | Problemas de acceso a internet | Utilización de proveedores de Internet alternativos como por ejemplo sistemas de Internet 4G que requieran de menor tiempo de instalación. En caso de situación crítica sin solución por parte del proveedor, se valoraría el cambio de ISP. |
| Riesgos de la organización | Fallo de hardware del equipo informático a utilizar | Obtención de un dispositivo de reserva y realización de copias de seguridad de los materiales del proyecto, tanto en dispositivos físicos externos al equipo como en servicios de almacenamiento en la nube |
| Riesgos de la dirección de proyectos | Estimaciones erróneas de la duración de las tareas | Reajuste de los plazos de realización de las tareas |

## LEGISLACIÓN Y NORMATIVA

En este apartado se hará una enumeración y revisión de las leyes y normas aplicables al proyecto a realizar. Se incluyen tanto aquellas de ámbito nacional como las que pertenezcan al ámbito europeo.

* **Reglamento General de Protección de Datos de la Unión Europea** [X]: conocido también como RGPD, establece los parámetros a la hora de realizar el tratamiento de datos personales de personas físicas y de su libre circulación. En este caso, como se verá más tarde durante el análisis, el malware realiza una recopilación indiscriminada de datos personales que se encuentran almacenados en el dispositivo infectado. Estas operaciones se realizan sin el conocimiento y autorización del usuario del dispositivo, y por supuesto también sin ninguna base legal para el tratamiento. Asimismo, estos datos son posteriormente exfiltrados a servidores que aparentemente estarían ubicados en un país fuera de la Unión Europea, y que podrían pertenecer a la categoría de “Países terceros” contemplada por el reglamento, lo cual estaría prohibido de no estar aplicadas las salvaguardas necesarias.
* **Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales** (Ley Orgánica 3/2018) [X]: conocida también como LOPD-GDD, cuya función es la adaptación del marco legal español al RGPD dictado por la Unión Europea. Nuevamente, el malware a analizar no cumple los requisitos que establece la ley a la hora de obtener y extraer del dispositivo infectado los datos de carácter personal. Estos datos recopilados no se incluyen dentro de las naturalezas o índoles que están exentas de base legitimadora y la aplicación de tipo troyano analizada carece de los permisos necesarios por parte de los usuarios, y tampoco otorga a estos los derechos necesarios respecto a sus datos establecidos por la ley.
* **Convenio de Budapest** [X]: tratado a nivel europeo cuyo fin es combatir la ciberdelincuencia mediante entre otras medidas la colaboración entre diferentes instituciones pese a que estas pertenezcan a diferentes países, así como tratar de que sus respectivas leyes estén armonizadas en lo respectivo a los delitos informáticos. Este convenio recoge las acciones realizadas por el malware a analizar como delictivas, ya que se trata de una utilización de los recursos del dispositivo infectado, junto con una recopilación y posterior extracción de datos de este que en ningún momento ha sido autorizada por el usuario y está siendo realizada de manera ilegal.
* **Plan Nacional de Ciberseguridad** y **Real Decreto 311/2022** [X]: real decreto cuyo objetivo es la regulación del Esquema Nacional de Seguridad para determinar la política de seguridad en la utilización de los medios electrónicos y adaptarlo a la transformación digital de la sociedad. En el caso de que el malware bajo análisis afectase a dispositivos que formasen parte del sector privado y se produjese una brecha de datos, se debería de seguir el procedimiento de notificación establecido y mediante su presentación ante el organismo INCIBE-CERT.
* **INCIBE-CERT** [X]: organismo de respuesta a incidentes del Instituto Nacional de Ciberseguridad. Dentro de su Guía Nacional de Notificación y Gestión de Ciberincidentes [X] se establece un marco de referencia que ha sido consensuado entre los organismos nacionales competentes en el ámbito de la notificación y gestión de incidentes. Exige unos criterios mínimos y unas obligaciones de reporte para los casos de incidente recogidos por la ley. Dependiendo del tipo de incidente provocado por el malware bajo análisis, este caso podría producirse y por lo tanto, se deberían seguir la metodología establecida para llevar a cabo tanto la notificación del incidente como el seguimiento, todo ello mediante el sistema de ventanilla única. En función del tipo de afectado por el malware también se podrían aplicar requerimientos particulares por parte de las autoridades competentes.

**Capítulo 3: Solución**

A lo largo de este capítulo se presentará la solución al problema del proyecto. Se incluirá una descripción general sobre esta, además de detallar el proceso que se ha seguido para su obtención y los resultados finales obtenidos.

## 3.1 DESCRIPCIÓN DE LA SOLUCIÓN

Para comprender el funcionamiento del malware se deben realizar los análisis planificados. En la fase de análisis estático no se ejecuta el malware, si no que se obtienen en primer lugar marcas identificadoras del malware y se realiza posteriormente un análisis del código de la aplicación, bien sea en los archivos DEX presentes, o analizando el código Java o Kotlin disponible. Es posible que se deba realizar una tarea de deofuscación previa en el caso de no encontrarse el código inicial en un estado legible, lo cual suele ser común en aplicaciones malware que tratan de esconder sus intenciones maliciosas de los analistas.

En la sección de análisis dinámico se ejecutará el malware utilizando el entorno controlado con dispositivos emulados mencionado en apartados anteriores. Se llevarán a cabo numerosas pruebas para tratar de cubrir los caminos de ejecución descubiertos durante el análisis estático. Además, en el caso de contar con servidores C&C, se analizará la comunicación del troyano con estos mediante un sistema de proxy. Si por el contrario los servidores externos no estuviesen accesibles o no respondiesen como se esperase, impidiendo por lo tanto la exploración de nuevos caminos de ejecución, se procedería a tratar de interceptar el tráfico desde el dispositivo infectado y tratar de redirigirlo a un servidor local en el que se intentaría simular el servidor externo y con el cual se podría tratar de provocar comportamientos del malware determinados observados previamente en el análisis estático.

Una vez finalicen ambas fases, esto es, cuando se hayan recabado los datos necesarios para determinar los comportamientos y capacidades que posee el troyano, se procede a la comparación de resultados entre las dos fases, así como a la presentación de toda la información que se considere relevante en el posterior informe.

## 3.2 EL PROCESO DE DESARROLLO

A continuación, se detalla el proceso que se ha seguido en las fases de análisis del malware. En primer lugar, se llevará a cabo una identificación de la muestra de la familia Chameleon elegida junto con un análisis automático de esta. Posteriormente se continuará con las dos fases de análisis, que como se ha mencionado previamente, se llevan a cabo de manera simultánea para tratar de extrapolar resultados obtenidos y hacer que los datos obtenidos en una fase sirvan para progresar en la otra y viceversa.

### 3.2.1 Identificación de muestra

Se realizará una obtención de valor hash con la función SHA-256. De esta manera, se obtendrá un identificador único para la muestra.

Además de este identificador, se hará un escaneo automatizado utilizando la herramienta Apkid. El propósito es observar qué tipo de protecciones puede tener el malware de cara al análisis. Si se han utilizado herramientas conocidas y populares para el packing o la ofuscación del troyano, estas serán detectadas, y se podrá proceder a investigar sobre procesos específicos para intentar revertir estas medidas anti-análisis.

En resumen, las subtareas asociadas con esta primera fase de la solución son las siguientes:

* Obtención de identificador hash único.
* Identificación automática de posibles métodos anti-análisis con Apkid.

### 3.2.2 Análisis estático

En esta fase se llevará un análisis del código de la muestra APK sin realizar su ejecución. En primer lugar, se procederá a desempaquetar el fichero de la muestra y hacer una enumeración de sus contenidos más importantes. Algunos de los archivos adicionales, como puede ser el manifiesto de Android o los recursos, pueden ser de gran ayuda a la hora de identificar posibles comportamientos del malware previa a su ejecución, además de poder arrojar más detalles sobre posibles métodos de ofuscación y anti-análisis.

En la fase de análisis puro se utilizará JADX para obtener código Java a partir del fichero DEX de la aplicación, ya que este es más legible y fácil de comprender que el bytecode DEX de ART. Una vez obtenido el código Java, se procederá a analizar sus clases e intentar comprender comportamientos y funciones que es probable sean efectuadas por el malware en el momento de su ejecución. En el caso de contar con código ofuscado, se procederá a su renombrado e inclusión de comentarios para facilitar la tarea de comprensión.

Por lo tanto, las actividades previstas inicialmente en la solución del análisis estático son las siguientes:

* Desempaquetado del archivo APK de la muestra.
* Enumeración de los archivos principales encontrados.
* Análisis de los archivos importantes:
  + Manifiesto XML Android y permisos asociados.
  + Recursos utilizados por la aplicación.
  + Archivos adicionales que pueden ser de utilidad.
* Conversión de código DEX a Java mediante JADX.
* Análisis del código Java encontrado.
* Anotación y renombrado en caso de ofuscación.

3.2.3 Análisis dinámico

Paralelamente al análisis estático, se instalará y ejecutará la muestra del troyano en los dispositivos Android emulados que proporcionan tanto Android Studio como Genymotion en el caso de necesitarse acceso de superusuario o la instrumentación mediante FRIDA para alterar el funcionamiento. Se recabará toda la información pertinente del proceso de instalación, como la petición de permisos o la generación de notificaciones o alertas, así como el comportamiento en términos generales del malware. Posteriormente, se analizará el comportamiento visible de la aplicación instalada en el dispositivo emulado. Simultáneamente, se llevará a cabo un análisis del tráfico generado y recibido por parte del dispositivo, tratando de interceptar y diseccionar los paquetes que hayan sido enviados por el malware en el caso de comunicarse con servidores C&C externos.

Un punto importante que remarcar es que en el supuesto caso de que los servidores remotos no se encuentren disponibles, lo cual es un hecho posible ya que el ciclo de vida de este tipo de malware con servidores C&C es relativamente corto [X], se puede intentar crear un servidor local que simule ser el servidor remoto. Para ello habría que crear una infraestructura que respondiese a las peticiones del dispositivo infectado y enviase respuestas de acuerdo con los protocolos establecidos. Además, se deberían de redirigir las peticiones del malware al nuevo servidor local.

Esta solución de servidor local podría ser utilizada no solamente en el caso de no estar disponibles los servidores remotos de los atacantes, sino también en el caso de querer observar algún comportamiento determinado del malware identificado en el análisis estático, pero no desencadenado por el servidor remoto. Desde el servidor local se podría controlar con mayor granularidad el comportamiento del malware, permitiendo la exploración de los caminos de ejecución encontrados sin dependencias externas.

Finalmente, las subtareas que pertenecen a esta fase dinámica de la solución son las siguientes:

* Instalación de la muestra en dispositivos emulados.
* Ejecución de la muestra.
* Monitorización del comportamiento visible provocado por la muestra.
* Captura y disección de tráfico de red generado por la muestra.
* Intento de creación de servidor local de simulación C&C.

Es importante recalcar que, debido a la naturaleza de gran parte de las tareas, un componente importante es el de la propia investigación. Además, no se conocen previamente las posibles medidas que pueda presentar el malware para impedir su análisis. Por lo tanto, parte de la solución propuesta podría verse modificada a la hora de su realización si fuese necesario debido a elementos que forzasen una alteración del proceso.

## 3.3 RESULTADOS DEL DESARROLLO

**Capítulo 4: Evaluación**

### 4.1 Identificación de la muestra

Se realiza en primer lugar el hashing utilizando SHA-256 para obtener el valor único que identifica a la muestra seleccionada. Como se ha comentado previamente, la muestra es una versión de la familia Chameleon que surgió por primera vez en Noviembre de 2023. Se ha obtenido del repositorio VX-Underground, y ha sido también la muestra utilizada por la empresa Cyble en el único análisis conocido del malware, pese a que el informe que han realizado no se encuentra disponible públicamente. Es una muestra que se hace pasar por el popular navegador web Google Chrome. Se trata por lo tanto de un troyano de la categoría de aplicaciones conocidas, tratando de utilizar la popularidad de Chrome para aumentar su número de descargas y simulando ser una aplicación conocida y benigna.

El resultado de la función hash devuelve 0a6ffd4163cd96d7d262be5ae7fa5cfc3affbea822d122c0803379d78431e5f6 como el identificador único del archivo APK obtenido.

A continuación, utilizaremos Apkid para realizar un análisis básico automatizado y observar si el malware utiliza algún método estándar de packing u ofuscación, ya que este dato sería de extrema utilidad para las próximas fases. Tras utilizar Apkid, obtenemos los siguientes resultados:

[+] APKiD 2.1.5

[\*]0a6ffd4163cd96d7d262be5ae7fa5cfc3affbea822d122c0803379d78431e5f6.apk!classes.dex

|-> anti\_vm : Build.MANUFACTURER check

|-> compiler : dexlib 2.x

Pese a utilizar el modo verboso de la herramienta, no extrae una gran cantidad de información de la muestra. No obstante, con esta información obtenemos dos datos que nos serán de utilidad. En primer lugar, es posible que el malware tenga un mecanismo de protección para prevenir ciertos comportamientos en el caso de encontrarse en un entorno emulado (comprobación anti\_vm). En segundo lugar, el compilador utilizado es dexlib en su versión 2. Esta no es una elección estándar [X], y se trata de un compilador común en casos en los que aplicaciones han sido inyectadas con malware previo a su reempaquetado. Esto encaja con la naturaleza de la aplicación al ser un troyano que se hace pasar por Google Chrome.

### 4.1.2 Análisis estático

En esta fase se comenzará tratando de cargar el APK en Android Studio o JADX para intentar observar su código fuente y comenzar el análisis. Tras crear un nuevo proyecto en Android Studio para el APK, observamos tres paquetes diferentes en el archivo classes.dex con subpaquetes en dos de los casos, siendo la siguiente la estructura jerárquica inicial del proyecto:

* Com.busy
  + Gather
  + Lady
  + Private
  + Soft
  + Stem
  + Unaware
  + wedding
* Com.microsoft.office
  + Officespace
  + Orapi
  + Otcui
  + Plat
  + playStoreDownloadManager
  + privacy
  + process
  + resourcedownloader
  + resources
* Com.yandex
  + Metrica

Dados los nombres del primer paquete y sus subpaquetes, parecen haber sido ofuscado, o al menos no estar organizado de una manera lógica para ser legible. Los dos siguientes paquetes parecen hacer referencia a la suite Office de Microsoft y al sistema de recolección de métricas de Yandex, el buscador web ruso más popular. No obstante, habrá que comprobar el código presente en ellos para cerciorarse de sus funcionalidades y ver si realmente hacen referencia a estos servicios.

Otro archivo que es importante analizar inicialmente para comprender las posibles funcionalidades del malware es el manifiesto de permisos. Este archivo está presente en todas las aplicaciones de Android, y es la declaración de la aplicación al sistema sobre los tipos de permisos que va a solicitar. Siempre se encuentra en la raíz del archivo APK desempaquetado con el nombre “AndroidManifest.xml”. Utilizaremos Apktool para obtener desempaquetados los archivos presentes en el fichero APK inicial, y los almacenaremos en el directorio “muestra\_empaquetada”. Se presenta a continuación la tabla con las entradas encontradas en el XML de permisos junto con el tipo de accesos que otorga cada permiso de acuerdo con la documentación proporcionada por Google [X]:

|  |  |
| --- | --- |
| **Nombre del permiso** | **Acciones permitidas** |
| WRITE\_EXTERNAL\_STORAGE | Permite la escritura a almacenamiento externo |
| READ\_EXTERNAL\_STORAGE | Permite la lectura de almacenamiento externo |
| REQUEST\_IGNORE\_BATTERY\_OPTIMIZATIONS | Permite solicitar al usuario la opción de ignorar optimizaciones de batería para la aplicación |
| FOREGROUND\_SERVICE | Permite el comienzo de un servicio en primer plano |
| REQUEST\_INSTALL\_PACKAGES | Permite que la aplicación solicite instalar paquetes |
| GET\_PACKAGE\_SIZE | Permite conocer el tamaño de cualquier paquete |
| PACKAGE\_USAGE\_STATS | Permite la recopilación de estadísticas de uso de paquetes |
| ACCESS\_BACKGROUND\_LOCATION | Permite el acceso a la ubicación cuando la aplicación se encuentra en segundo plano |
| ACCESS\_FINE\_LOCATION | Permite el acceso a la ubicación precisa |
| ACCESS\_COARSE\_LOCATION | Permite el acceso a la ubicación aproximada |
| WRITE\_SECURE\_SETTINGS | Permite la lectura y escritura de ajustes seguros del sistema |
| GET\_TASKS | Permite acceso a tareas (valor deprecado desde Android 5.0) |
| DISABLE\_KEYGUARD | Permite desactivar el keyguard (pantalla de bloqueo) en el caso de no considerarse seguro |
| RECEIVE\_LAUNCH\_BROADCASTS | Función no recogida en la documentación de Android |
| READ\_CONTACTS | Permite la lectura de los contactos del usuario del dispositivo |
| RECEIVE\_BOOT\_COMPLETED | Permite la recepción de una transmisión notificadora cuando el dispositivo finaliza su arranque |
| WRITE\_CALL\_LOG | Permite la escritura en el registro de llamadas del usuario |
| READ\_CALL\_LOG | Permite la lectura en el registro de llamadas del usuario |
| QUERY\_ALL\_PACKAGES | Permite la obtención de información de todos los paquetes presentes en el dispositivo |
| ACTION\_MANAGE\_OVERLAY\_PERMISSION | Permite la superposición de gráficos sobre otras aplicaciones |
| ACCESS\_WIFI\_STATE | Permite el acceso a la información sobre redes Wi-Fi |
| CHANGE\_WIFI\_STATE | Permite modificar el estado de conexión Wi-Fi |
| BLUETOOTH\_SCAN | Permite el escaneo de dispositivos Bluetooth |
| READ\_PHONE\_STATE | Permite la lectura de información de red telefónica y el estatus de llamadas activas |
| POST\_NOTIFICATIONS | Permite la creación de notificaciones |
| WRITE\_CONTACTS | Permite la escritura de información en los contactos del usuario |
| BLUETOOTH | Permite la conexión a dispositivos Bluetooth cercanos |
| MODIFY\_PHONE\_STATE | Permite realizar cambios al estado de telefonía del dispositivo (desactivación, cambios en el código MMI) |
| CAMERA | Permite el acceso a la cámara del dispositivo |
| SEND\_SMS | Permite el envío de mensajes SMS |
| RECEIVE\_SMS | Permite la recepción de mensajes SMS |
| READ\_SMS | Permite la lectura de mensajes SMS |
| READ\_PHONE\_NUMBERS | Permite el acceso a los números de teléfono del dispositivo |
| WAKE\_LOCK | Permite el acceso al PowerManager del sistema, encargado de mantener el procesador activo o la pantalla encendida |
| LOCK\_DEVICE | Permite el bloqueo del dispositivo |
| INTERNET | Permite la apertura de sockets de red por parte de la aplicación |
| CALL\_PHONE | Permite el inicio de una llamada telefónica sin pasar por la interfaz de llamada estándar |
| SYSTEM\_ALERT\_WINDOW | Permite la creación de superposiciones en todas las aplicaciones con el modelo de interacciones utilizado por las interacciones del sistema |

Podemos observar que la aplicación solicita una gran cantidad de permisos que pertenecen a categorías muy diferentes, además de algunos que están categorizados como permisos con nivel de peligro en cuanto a su protección por parte de las recomendaciones de Google, como ACCESS\_FINE\_LOCATION, READ\_SMS o CALL\_PHONE.

Además de la información de permisos, del manifiesto XML también podemos encontrar el punto de entrada de la aplicación, que en Android se indica con la acción principal ("android.intent.action.MAIN") y la categoría LAUNCHER. En este caso existen múltiples puntos de entrada registrados, que hacen referencia a diferentes aplicaciones populares como YouTube, Gmail o Telegram. Sin embargo, todas estas entradas están desactivadas menos una, que carga dinámicamente el nombre de la aplicación, y registra la actividad de la clase “Xk011b72bcbf5396b4f9ec9737b706f44.Xkfb2bcfa7ae858b5247a4fc3c5ae3020.Xk1b4219fb6ddeebbcfdc5d80780f8b2f.Xk7d4258c1c9c337733fb3a8c3aaf570e” como la inicial. Dado el nombre de la clase declarado, parece que el código de la aplicación podría estar ofuscado o que realiza una carga dinámica de otras clases durante la ejecución.

Por último, el manifiesto también arroja información sobre las intenciones de la aplicación respecto a los permisos de accesibilidad y la API de administración de servicios entre otros.

Para tener acceso a estas funciones, la aplicación declara en el manifiesto una serie de servicios y receptores. Los servicios son los siguientes:

* Un servicio de enlace con el permiso BIND\_ACCESSIBILITY\_SERVICE. Este permiso permite la recepción y emisión de eventos de accesibilidad del sistema [X]. Este tipo de eventos va desde acciones de clic en la pantalla, deslizamiento o clic en los botones de navegación del sistema. A través de este permiso se puede tener una perspectiva general de qué acciones se realizan por parte del usuario, además de poder realizar de manera automática las acciones deseadas.
* Un servicio de enlace con el permiso BIND\_NOTIFICATION\_LISTENER\_SERVICE [X]. En este caso el permiso proporciona en forma de eventos todas las notificaciones que se producen en el sistema Android, además de la reacción de los usuarios a estas (cancelaciones de la notificación, notificaciones descartadas por tiempo de espera excesivo…).
* Un servicio de enlace con el permiso SEND\_RESPOND\_VIA\_MESSAGE, que permite a una aplicación responder a una llamada a través de un mensaje. De acuerdo con la documentación de Android, este permiso no debería ser solicitado por aplicaciones de terceros, estando reservado para aplicaciones del sistema.
* Una serie de servicios que están asociados con el permiso BIND\_JOB\_SERVICE, utilizado para la gestión de peticiones asíncronas. Todos estos servicios tienen seleccionado como proceso a “:acra”. ACRA es un conocido sistema de informe de fallos (más conocido como crash reporting) de código libre, utilizado en una gran cantidad de aplicaciones Android [X]. Este conjunto de servicios podría indicar la intención de la aplicación de llevar a cabo peticiones con información de fallos en el caso de producirse problemas de ejecución del malware.

Como último dato relevante del manifiesto, podemos observar datos adicionales relativos a la compilación de la aplicación, ya que la versión mínima del SDK Android que soporta es la 22, que hace referencia a Android 5.1 (conocido como Lollipop), mientras que la versión SDK objetivo para la que se ha desarrollado la aplicación es la 26, que se trata de Android 8.0 (conocido como Oreo).

Una vez extraída esta información a partir del manifiesto XML, podemos pasar al resto de archivos resultantes del desempaquetado del archivo APK. Un archivo clave será el ejecutable DEX mencionado previamente, que contiene información acerca de las clases iniciales que componen la aplicación, y a partir del cual podemos tratar de obtener código en formato Java. Para ello utilizaremos la herramienta JADX en su versión con interfaz gráfica de usuario.

Una vez JADX descompila el DEX a código Java, podemos navegar por las clases encontradas como si fuese un proyecto Android estándar:



Sin embargo, parece que todo el código de los subpaquetes de com.busy se encuentra ofuscado, con nombres de métodos que no son relevantes y código deliberadamente complejo para tratar de esconder su funcionamiento. Además, no está presente la clase de la actividad inicial registrada en el manifiesto XML.



Por lo tanto, debemos asumir que la aplicación ha sido ofuscada y que posiblemente realice carga dinámica de clases para evitar que estas sean analizadas de manera sencilla. Respecto a la ofuscación, ya que Apkid no ha detectado ningún método de los más comunes como ProGuard o DexGuard, no poseemos mecanismos automáticos para tratar de revertir el proceso, por lo tanto, se procederá a investigar si se realiza una carga dinámica de clases. En el caso de realizarse, Android siempre utiliza la clase DexClassLoader [X] para la carga adicional desde un archivo JAR o APK. Encontramos la creación de un objeto DexClassLoader en la clase “YRoEyIoHcCt” mostrada anteriormente.

En este caso, la pila de llamadas es excesivamente compleja e incluye multitud de operaciones intermedias, con lo que impide averiguar el primer argumento pasado al constructor de DexClassLoader, que es el directorio del archivo DEX que utilizará para cargar las nuevas clases. Por ello, vamos a ayudarnos de la información extraída en la fase de análisis dinámico y utilizando el depurador de Android Studio trataremos de observar los parámetros utilizados para la carga dinámica.

Para finalizar esta fase de análisis estático, continuaremos con el análisis de los archivos desempaquetados con el fin de intentar obtener más información acerca del malware. En la carpeta de recursos estándar de Android (nombrada “assets”), podemos encontrar varios ficheros que podrían resultar útiles:

* Se incluyen cuatro certificados con el prefijo “AmazonRootCA”. Tras tratar de obtener información relativa al emisor del certificado mediante OpenSSL (openssl x509 -noout -text -in AmazonRootCA1.cer), tres de ellos no pueden ser leídos correctamente sin ningún fallo específico, mientras que el cuarto certificado presenta errores de estructura que provocan que tampoco se puede cargar.
* Archivos multimedia de sonido (formato OGG) que contienen tonos de notificación y sonidos de funciones del sistema.
* Archivos de imagen PNG que contienen parte de la firma hexadecimal de formato PNG (89 50 4E) tras su lectura con la herramienta HexFiend, pero que no pueden ser leídos correctamente por los programas de visualización estándar del sistema operativo.
* Un archivo de texto “hello.txt” que está codificado en base64. Tras decodificarlo utilizando la terminal de comandos de Mac (cat hello.txt | base64 -D). El texto resultante parece estar relacionado con la obtención de permisos, contando con la traducción a múltiples idiomas de numerosos mensajes relativos al proceso de permisos, a continuación, se muestra un extracto de los mensajes en español del fichero:
  + "permission\_access\_only\_foreground":["Permitir solo con la app en uso"]  
    "allow":["Permitir"]  
    "next\_label":["Siguiente"]  
    "application\_info\_label":["Información de la aplicación", "Info. de las apps"]  
    "service\_stop":["Detener"]  
    "remove\_device\_admin":["Desactivar","Desactivar este administrador de dispositivos"  
    "Desactivar esta app de administración del dispositivo"]  
    "add\_device\_admin\_msg":["¿Deseas activar el administrador del dispositivo?"  
    "¿Activar administrador de dispositivos?"]
* Varios ficheros JSON haciendo referencia a gráficos vectorizados de Adobe y a algunos de los archivos de imagen que no han podido ser leídos correctamente previamente, con lo que es posible también que se realicen modificaciones en tiempo de ejecución a los archivos multimedia para ocultarlos del análisis estático.
* Un archivo en formato JSON que contiene información relativa a varias regiones y ciudades de China, incluyendo un sistema de codificación y URLs apuntando a archivos ZIP. Las entradas de la lista de ciudades contienen una propiedad llamada “adcode”, lo que podría indicar un sistema de publicidad con diferentes zonas.
* Un archivo con extensión JSON (kx.json), que es el que más tamaño tiene (815KB), pero que parece no ser realmente un fichero JSON, ya que no puede ser leído como tal. No obstante, el formato JSON no tiene bytes reconocibles que permitan extraer una firma de formato [X]. Sin embargo, sí que podemos observar la representación hexadecimal del archivo, utilizando nuevamente la herramienta Hex Fiend:
  + Obtenemos los primeros bytes hexadecimales: 50 4B 03 04. Esta secuencia hace referencia al formato comprimido ZIP. Además, en la representación de texto plano del fichero observamos en la primera línea una referencia a classes.dex, por lo tanto, este podría tal vez ser el archivo que se carga dinámicamente mediante DexClassLoader:



Para comprobar si efectivamente el archivo “kx.json” contiene un archivo DEX de carga dinámica en su interior, procedemos a tratar de abrirlo como un fichero ZIP, y confirmamos que se descomprime y se obtiene un nuevo archivo “classes.dex” que aumenta en tamaño hasta los 2,2MB. Iniciamos nuevamente el proceso de descompilación de DEX a código Java con JADX con el nuevo archivo DEX encontrado. Obtenemos nuevamente código ofuscado, pero en este caso más legible que en la primera iteración. Además, encontramos la clase de tipo Activity marcada como punto de entrada de la aplicación en “AndroidManifest.xml”, que tiene el nombre “Xk7d4258c1c9c337733fb3a8c3aaf570e”.

Por lo tanto, hasta el momento nos encontraríamos con la siguiente estructura que sugiere el código en lo referente a la carga de clases de manera dinámica:



Como se ha mencionado, las nuevas clases obtenidas también se encuentran ofuscadas, pero su legibilidad es mayor y al contar con el punto de entrada de la aplicación, se puede tratar de realizar una deofuscación manual mediante las funciones que posee JADX para el renombrado de clases y métodos, además de su sistema de comentarios de código con el que dejar referencias para futura investigación.

En la siguiente parte del análisis estático, se irán mostrando las clases diseccionadas durante el proceso. No se incluyen todas las clases pertenecientes a la aplicación, ya que algunas de ellas no son pertinentes, y en el caso de otras no se ha podido confirmar que lleven a cabo alguna función dentro de los mecanismos que exhibe el malware.

Las clases presentadas han sido renombradas respecto a sus versiones ofuscadas para mejorar la legibilidad y hacer el sistema más fácil de comprender.

En primer lugar, se tratarán las clases descubiertas relativas al punto de entrada de la aplicación y el mecanismo que utiliza para la obtención de los permisos necesarios, así como las funciones de la API de administración y de accesibilidad comentadas en capítulos anteriores. Las principales clases involucradas en los procesos mencionados se encuentran en el siguiente diagrama UML:



Durante el proceso de intento de obtención de permisos también se realizan varias comprobaciones sobre el dispositivo infectado, alterando los mecanismos utilizados en función tanto de la versión de Android encontrada como de los resultados de las verificaciones de fabricante. Para realizar estas validaciones de marca, Chameleon hace uso de las siguientes clases:

* DeviceActionManager
* AsusChecker
* HtcChecker
* HuaweiChecker
* LetvChecker
* MeizuChecker
* MeizuSecurityVersionEnum
* OnePlusChecker
* SamsungChecker
* XiaomiChecker
* ZteCheker
* DeviceManufacturersEnum
* DeviceBuildDetailsFinder

El malware trata de almacenar información variada, incluyendo el método de desbloqueo del dispositivo en el caso de que este sea el método elegido por el usuario. Además, introduce la función de keylogger. Este mecanismo es utilizado en este caso para registrar la contraseña o el PIN de desbloqueo elegido por el usuario, pero también se utilizará en otras funciones de recogida de datos de tipología infostealer.



A continuación, se lleva a cabo una revisión de las clases encontradas relativas a la gestión de la base de datos creada por el malware y utilizada para almacenar datos generados por el dispositivo infectado con la intención de su posterior exfiltración a los servidores C&C. El diagrama de clases involucradas en estas tareas es el siguiente:



Posteriormente, se presentan las clases relativas a la función de exfiltración de datos que posee el malware:



Como se puede observar, se realizan distintos tipos de exfiltración en función de las condiciones que se presenten y los comandos que reciba el malware. Estos comandos, que serán comentados en más detalle posteriormente, también permiten la instalación de nuevas aplicaciones a partir de ficheros en formato JAR.

Esta instalación de nuevas aplicaciones también está asistida por la clase ModuleManager:

Chameleon también realiza un análisis de las aplicaciones instaladas en el dispositivo, creando una serie de tareas en segundo plano que son las encargadas de recopilar la información acerca de las aplicaciones que será posteriormente exfiltrada y almacenarla en la base de datos. Las clases mostradas son las que gestionan el proceso:



Un punto importante es que el malware elige no recopilar información acerca de aplicaciones desarrolladas con React. Se desconoce la razón de este comportamiento ya que no existen más referencias en el código que permitan expandir el contexto. El siguiente código desofuscado es la comprobación realizada que excluye a React:



Parte de los datos que extrae Chameleon son relativos a las posibles cookies almacenadas en el navegador web sistema, y las siguientes clases son las encargadas de gestionar la recopilación y el procesado de esta información:

* CookieManager
* CookieManagerExtended

Una de las funciones del troyano es la recepción de diferentes tipos de eventos y transmisiones que se producen en el dispositivo infectado. En función del tipo de acción ocurrida y sus características específicas, el malware puede tomar distintos caminos. Para tener constancia de todos estos eventos, la aplicación registra la siguiente lista de clases:

* ScreenBroadcastReceiver
* PackageChangesBroadcastReceiver
* CloseSysDialogBroadcastReceiver
* RiseUpBroadcastReceiver
* WakeLockBroadcastReceiver
* PreferenceFileListener
* NotificationListenerServiceImpl
* SensorManager

Otra funcionalidad que presenta el malware es la gestión de SMS, incluyendo la lectura de su contenido y también su borrado si el mensaje recibido cumple con unas características específicas. Las siguientes clases son las utilizadas para realizar estas tareas:

* SmsBroadcastReceiver
* SmsManager
* SmsRemover

Para gestionar las comunicaciones con los servidores C&C remotos, el malware hace uso de las siguientes clases:



Un aspecto destacable de esta funcionalidad es la resiliencia del sistema en el caso de errores en los servidores remotos. Además de realizar comprobaciones periódicas y almacenar la información de la última conexión en la base de datos, el malware reacciona en caso de que los servidores remotos hayan sufrido problemas, y en el caso de recibir un status de denegación a la petición enviada, procede a eliminar toda la información acerca de los servidores de la base de datos local, suprimiendo posibles trazas que pudiesen ser analizadas a posteriori.

Uno de los procesos claves que guía el comportamiento del troyano es el sistema de preferencias. Se encuentra almacenado en el archivo com.busy.lady\_preferences.xml dentro del directorio de datos del dispositivo infectado. La estructura de datos utilizada es un mapa XML que cuenta con el nombre de la preferencia como valor llave. No obstante, todos los valores del mapa se encuentran encriptados, con lo que inicialmente no se pudo averiguar qué tipo de ajustes servían para controlar el sistema.

Durante la realización del análisis estático, se encontraron las siguientes clases a cargo de la encriptación y desencriptación de valores en el sistema. En primer lugar, una clase que se ha renombrado a “EncryptionService”, que tiene como función la encriptación y codificación de los valores de preferencias mencionadas anteriormente.



Posteriormente, se encontró la clase para la operación contraria, la desencriptación de valores:



Tras tratar de aplicar ingeniería inversa al método de encriptación, se descubrió que la clave utilizada es el identificador único del dispositivo infectado mencionado previamente. Una vez conocida la clave y el vector de inicialización (que se encuentra directamente en el código y no sufre alteraciones), se puede proceder a la creación de una clase en Java que desencripte automáticamente los valores de los ajustes, llamada “ChameleonPreferenceDecryptor”. Hay que tener en cuenta el método utilizado para la encriptación inicial, que se trata del algoritmo AES en modo “block” y con un modo de relleno o padding conocido como PKCS5Padding [X]. Por último, el algoritmo utilizado realiza varias codificaciones previamente a la encriptación final.

El diagrama a continuación muestra los pasos del proceso de desencriptación empleado en el fichero Java para uno de los valores reales encontrados en el XML de preferencias:



Como se muestra en el diagrama, los valores finales obtenidos tras aplicar el algoritmo de desencriptación son legibles e indican la naturaleza de los comandos y ajustes locales con los que cuenta Chameleon.

Una vez comprobado que el algoritmo generado funciona correctamente, se puede aplicar a todos los ajustes iniciales obtenidos a partir del fichero obtenido a partir de la ejecución inicial del malware durante la fase de análisis dinámico, obteniendo la siguiente lista de ajustes junto con sus valores iniciales:

|  |  |
| --- | --- |
| **Nombre del ajuste** | **Valor** |
| launch\_time | 1717806586 |
| self\_defense | false |
| is\_keylogger | true |
| auto\_delete | false |
| push\_write | false |
| auto\_click | true |
| type | web |
| annoy | true |
| intercept\_cookies | false |
| interrupt\_biometric | true |
| first\_launch | true |
| hide\_after\_delay | true |
| injection | true |
| text | Play protect is working |
| android\_13 | true |
| service\_cham | false |
| is\_chameleon | false |
| Interval | 2000 |
| play\_protect | -1 |
| lockscreen\_grabber | true |
| connect\_server\_a11y\_granted | false |
| self\_notification\_hide | false |
| title | Google Play |
| subtext | (sin valor) |
| injection\_type | a11y |
| is | false |
| work\_type | 1 |
| self\_defense\_shutdown | false |

Como funcionalidad adicional, el malware cuenta con una clase dedicada al modo de defensa del malware en el caso de que el comando remoto para la acción sea recibido, SelfDefenseManager.

Por último, el servicio de informe de fallos mencionado previamente, ACRA, cuenta con una serie de clases asociadas para gestionar detalles que se incluyen en los informes a enviar como el formato o qué tipo de datos se agregan:

* ACRAReportingManager
* ReportingService
* ReportFormatSpecifier
* AcraReportFieldsEnum

### 4.1.3 Análisis dinámico

**Capítulo 5: Conclusión**

En este apartado se comentarán los resultados obtenidos previamente, las aportaciones que se han realizado, se mencionan posibles ampliaciones que se podrían realizar en el futuro, se mencionan los problemas encontrados durante la realización de la tesis y, por último, se encuentra una sección de opiniones en la que se emitirán algunas valoraciones a título personal tras la realización del proyecto.

## 5.1 APORTACIONES REALIZADAS

## 5.2 AMPLIACIONES FUTURAS

## 5.3 PROBLEMAS ENCONTRADOS

## 5.4 OPINIONES PERSONALES

**Bibliografía**

|  |  |
| --- | --- |
| [1] | <https://www.researchgate.net/publication/312129902_Android_Fragmentation_Classification_Causes_Problems_and_Solutions> |
| [2] | <https://www.cs.tufts.edu/comp/116/archive/fall2017/tzhu.pdf>) |
| [3]  [4]  [5]  [6] | <https://f-droid.org/es/>  <https://www.kaspersky.com/blog/malware-in-google-play-2023/49579/>  <https://developers.google.com/android/play-protect/potentially-harmful-applications>  <https://source.android.com/docs/core/permissions/privacy-indicators>) |

|  |  |
| --- | --- |
| [7]  [8]  [9]  [10]  [11]  [12]  [13]  [14]  [15]  [16]  [17]  [18]  [19]  [20]  [21]  [22]  [23]  [24]  [25]  [26]  [27]  [28]  [29]  [30]  [31]  [32]  [33]  [34]  [35]  [36]  [37]  [38]  [39]  [40]  [41] | <https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=en>  <https://stuff.mit.edu/afs/sipb/project/android/docs/guide/topics/admin/device-admin.html>  <https://cyble.com/blog/chameleon-a-new-android-malware-spotted-in-the-wild/>  <https://www.bbc.com/news/technology-21860360>  <https://www.europapress.es/portaltic/ciberseguridad/noticia-malware-chameleon-bloquea-huella-dactilar-forzar-introduccion-pin-hacerse-cuentas-bancarias-20231229135249.html>  <https://www.xatakandroid.com/seguridad/asi-logra-saltarse-biometria-movil-peligroso-troyano-bancario-chameleon>  <https://newsukraine.rbc.ua/news/android-virus-chameleon-banking-learns-to-1703623824.html>  <https://www.ndss-symposium.org/wp-content/uploads/2018/03/NDSS2018_04A-4_Duan_Slides.pdf>  <https://www.researchgate.net/publication/308180878_Structural_analysis_of_packing_schemes_for_extracting_hidden_codes_in_mobile_malware>  <https://developer.android.com/build/shrink-code#obfuscate>  <https://github.com/CalebFenton/simplify>  <https://media.defcon.org/DEF%20CON%2025/DEF%20CON%2025%20presentations/DEF%20CON%2025%20-%20Slava-Makkaveev-and-Avi-Bashan-Unboxing-Android.pdf>  <https://github.com/Guardsquare/proguard>  <https://media.defcon.org/DEF%20CON%2022/DEF%20CON%2022%20presentations/DEF%20CON%2022%20-%20Strazzere-and-Sawyer-Android-Hacker-Protection-Level-UPDATED.pdf>  <https://github.com/skylot/jadx>  <https://www.ghidra-sre.org/>  <https://www.researchgate.net/publication/262412421_DIVILAR_Diversifying_intermediate_language_for_anti-repackaging_on_Android_platform>  <https://kclpure.kcl.ac.uk/ws/portalfiles/portal/112357801/The_Evolution_of_Android_Malware_TAM_2017_GREEN_AAM.pdf>  <https://developer.android.com/studio/run/emulator>  <https://www.genymotion.com/>  <https://www.android-x86.org/>  <https://users.ece.cmu.edu/~tvidas/papers/SPSM14.pdf>  <https://www.cis.upenn.edu/~mhnaik/papers/fse12_full.pdf>  <https://en.wikipedia.org/wiki/List_of_file_signatures>  <https://www.seg-social.es/wps/portal/wss/internet/Trabajadores/CotizacionRecaudacionTrabajadores/36537>  <https://www.incibe.es/sites/default/files/contenidos/guias/doc/guia_nacional_notificacion_gestion_ciberincidentes.pdf>  https://www.recordedfuture.com/blog/2020-adversary-infrastructure-report  https://rednaga.io/2016/07/30/apkid\_andandroid\_compiler\_fingerprinting/  <https://metrika.yandex.ru/>  <https://developer.android.com/reference/android/Manifest.permission>  <https://developer.android.com/reference/android/accessibilityservice/AccessibilityService>  <https://developer.android.com/reference/android/service/notification/NotificationListenerService>  <https://www.acra.ch/>  <https://en.wikipedia.org/wiki/List_of_file_signatures>  <https://www.ibm.com/docs/en/zos/2.4.0?topic=rules-pkcs-padding-method> |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Anexo A: Control de versiones**

Se ha creado un [repositorio público en GitHub](https://github.com/raulng9/TFM_MalwareAnalysis) con los materiales utilizados en la realización del proyecto, incluyendo documentos con recursos utilizados, resultados derivados de los análisis, muestras de malware utilizadas y notas acerca del proceso.

**Anexo B: Seguimiento del proyecto**

## B.1 FORMA DE SEGUIMIENTO

A lo largo de todo el proceso se han realizado reuniones periódicas con el tutor para monitorizar el progreso llevado a cabo, así como dar recomendaciones y correcciones al alumno, además de tratar las próximas etapas a realizar.

## B.2 PLANIFICACIÓN INICIAL

## B.3 PLANIFICACIÓN FINAL