Ciberseguridad

Retos actuales y líneas de investigación de Inria

Inria | LIBRO BLANCO | N°03

Los libros blancos de Inria analizan los principales retos actuales en informática y matemáticas y muestran las acciones llevadas a cabo por nuestros equipos de proyecto ante estos retos. Su objetivo es describir el estado del arte de un tema determinado, mostrando su complejidad y presentar las direcciones de investigación existentes así como las emergentes y su previsto impacto social.
Este libro blanco ha sido editado por Steve Kremer, Ludovic Mé, Didier Rémy y Vincent Roca. Ellos han coordinado las contribuciones de los investigadores de los equipos de Inria.
Muchas gracias a Janet Bertot por la corrección de este documento, así como a François Pottier, Gabriel Scherrer y Benjamin Smith, que leyeron partes del mismo.

Del original:
Steve Kremer, Ludovic Mé, Didier Rémy, Vincent Roca. Cybersecurity: Current challenges and Inria's research directions. Inria, pp.172, 2019, Inria white book. hal-01993308
1º Edición
Fecha de publicación: Enero 2019, Francia

Traducción autorizada del idioma inglés de la edición publicada por Inria

Traducción al español realizada por Inria Chile

Coordinación general y revisión técnica: Nayat Sánchez Pi

Coordinación de producción: Julia Allirot y Katherine Lippi

Coordinación ejecutiva: Andrés Vignaga

Revisión gráfica y maquetación: Estudio Paretti y Katherine Lippi

Agradecimientos – Gracias a las siguientes personas de Inria Chile por sus aportes a la revisión técnica: Jaime Aranda, Sebastián Aranda, Hugo Carrillo, Hernán Lira, Luis Martí, Astrid Reyes, Luis Valenzuela, Natalia Vidal, Andrés Vignaga

Impreso por Lahosa
Impresión: Octubre 2022, Chile
ISBN: 978-956-09873-1-0

Impreso en Chile / Printed in Chile

EDITORES
Steve Kremer – Ludovic Mé – Didier Rémy – Vincent Roca.

LISTA DE COLABORADORES
Laurent Amsaleg – Nicolas Anciaux – Daniel Augot – Frédéric Besson – Nataliia Bielova – Luc Bouganim – Anne Canteaut – Claude Castelluccia – André Chailloux – Konstantinos Chatzikokolakis – Mathieu Cunche – Jérôme François – Teddy Furon – Georges Gonthier – Guillaume Gravier – Gilles Guette – Hélène Kirchner – Jean–Louis Lanet – Cédric Lauradoux – Arnaud Legout – Gérard Le Lann – Daniel Le Métayer – Gaëtan Leurent – Anthony Leverrier – Stéphane Mocanu – Christine Morin – Maria Naya Plasencia – Catuscia Palamidessi – David Pointcheval – Tamara Rezk – Michaël Rusinowitch – Kavé Salamatian – Guillaume Scerri – Nicolas Sendrier – Olivier Sentieys – Éric Totel Valérie – Viet Triem Tong.

LISTA DE PERSONAS QUE HAN APORTADO OBSERVACIONES ÚTILES
Gildas Avoine – Emmanuel Baccelli – Hugues Berry – Karthikeyan Bhargavan – Bruno Blanchet – Bertrand Braunschweig – Isabelle Chrisment – Hervé Debar – Claude Kirchner – Philippe Pucheral – Emmanuel Thomé – Frédéric Tronel – Damien Vergnaud – Emmanuel Vincent.

Resumen ejecutivo

Wikipedia define la ciberseguridad como "la protección de los sistemas informáticos contra el robo o daño de su hardware, software o información, así como contra la interrupción o el desvío de los servicios que prestan". Más concretamente, la ciberseguridad consiste en garantizar tres propiedades de la información, los servicios y la infraestructura TI: ***confidencialidad, integridad y disponibilidad.*** Por lo tanto, asegurar un sistema de información significa impedir que una entidad no autorizada acceda, altere o haga inaccesibles los datos informáticos, los servicios de computación o la infraestructura informática. Otra propiedad cada vez más importante es la privacidad, que puede considerarse como la confidencialidad del vínculo entre las personas y los datos. Hay que tener en cuenta que los términos seguridad y prevención se utilizan a veces de forma errónea. Mientras que la prevención se refiere a las amenazas accidentales, la seguridad se refiere a las amenazas intencionales. La seguridad y la prevención siguen siendo ámbitos muy distintos y bien identificados que se basan en hipótesis diferentes, y los mecanismos de protección contra las amenazas accidentales e intencionales suelen ser complementarios. En este libro blanco, limitamos nuestra atención a la seguridad.

La digitalización de nuestra sociedad está cambiando radicalmente la forma de utilizar los sistemas informáticos. Una gran parte de la población está continuamente conectada a Internet, utilizando un número asombroso de servicios diferentes. Al mismo tiempo, estamos permanentemente expuestos a ataques: nuestros datos sensibles pueden ser robados, modificados o destruidos. También vivimos con el riesgo de filtrar por error y de forma irreversible nuestra información privada en Internet. Las empresas, los Estados y sus infraestructuras críticas, que hoy en día están interconectadas, también son vulnerables. Los daños económicos y sociales de los ciberataques que tienen éxito pueden ser cuantiosos. La ciberseguridad se ha convertido, pues, en una preocupación general para todos, ciudadanos, profesionales, políticos y, en general, para todos los responsables de la toma de decisiones.

Este libro ofrece una visión general de las áreas de investigación en ciberseguridad, acompañada por los aportes de los equipos de Inria. El primer paso en implementar la ciberseguridad es identificar las ***amenazas*** y definir el ***modelo del atacante*** correspondiente. Las amenazas, como el ***malware***, los daños físicos o la ingeniería social, pueden tener como objetivo el hardware, la red, el sistema operativo, las aplicaciones o los propios usuarios. Por esto hay que diseñar mecanismos de detección y protección para defenderse de estas amenazas. Uno de los mecanismos principales es la criptografía: las ***primitivas criptográficas*** pueden garantizar la confidencialidad e integridad de los datos.

Estas primitivas tienen que estar sometidas a un ***criptoanálisis*** continuo para garantizar el máximo nivel de seguridad. Sin embargo, las primitivas criptográficas no son suficientes para garantizar la seguridad de las comunicaciones y los servicios: esta tarea requiere el uso de los llamados ***protocolos criptográficos***, que implementan interacciones más ricas sobre las primitivas. Los protocolos criptográficos son sistemas distribuidos: garantizar que alcancen sus objetivos, incluso en presencia de un adversario activo, requiere el uso de técnicas de verificación formal, que han tenido un gran éxito en este campo.

Si bien las primitivas y protocolos criptográficos son bases fundamentales para la seguridad, se necesitan ***servicios de seguridad*** adicionales, como la autenticación y el control de acceso, para aplicar una política de seguridad. Estos servicios de seguridad, normalmente proporcionados por el sistema operativo o los dispositivos de red, pueden ser atacados y a veces eludidos. Por lo tanto, las actividades en el sistema de información deben ser revisadas para detectar cualquier infracción de la política de seguridad. Por último, dado que los ataques pueden propagarse con extrema rapidez, el sistema debe reaccionar automáticamente o al menos reconfigurarse para evitar la propagación de los ataques.

Como se ha señalado anteriormente, la privacidad se ha convertido en una parte intrínseca de la ciberseguridad. Sin embargo, aunque a menudo se basa en primitivas y protocolos criptográficos, también tiene sus propias propiedades, técnicas y metodología. Además, el estudio de la privacidad requiere a menudo tener en cuenta aspectos legales, económicos y sociológicos.

Todos estos mecanismos de seguridad deben integrarse cuidadosamente en las aplicaciones de seguridad crítica. Esto incluye las aplicaciones tradicionales de prevención crítica, cada vez más conectadas y, por lo tanto, más vulnerables a los ataques a la seguridad, así como las nuevas infraestructuras que se ejecutan en la nube o están conectadas bajo el principio conocido como el Internet de las Cosas (IoT, por el inglés Internet of Things).

A pesar de los recientes y significativos avances en varias áreas de la ciberseguridad, todavía quedan abiertas importantes preguntas científicas. A continuación se presentan algunos retos seleccionados en los que Inria puede hacer nuevas e importantes contribuciones:

→ *Criptografía post-cuántica*. Se cree que la construcción de una computadora cuántica será factible en las próximas décadas y que la mayor parte de la criptografía que se utiliza hoy en día podría ser fácilmente descifrada con un ordenador de este tipo. Por lo tanto, es importante pensar ahora en la criptografía resistente

a la computación cuántica, ya que la información cifrada hoy puede seguir siendo sensible cuando aparezcan las computadoras cuánticas.

→ *Computación sobre datos encriptados.* La necesidad de computar sobre datos encriptados ha surgido con la aparición de la nube y la computación externalizada. Este problema puede resolverse mediante las técnicas denominadas cifrado homomórfico y cifrado funcional. En 2009 se logró un avance teórico con el primer esquema de cifrado totalmente homomórfico, pero este esquema seguía siendo completamente impracticable debido a su escasa eficiencia computacional. Se ha avanzado mucho desde entonces, pero es necesario seguir investigando; cualquier avance técnico significativo puede ser rápidamente explotado como una ventaja económica.

→ *Protocolos criptográficos extremo a extremo formalmente verificados.* La seguridad de los protocolos criptográficos es extremadamente difícil de garantizar y el uso de métodos rigurosos y formales es una necesidad. Las pruebas de seguridad asistidas por computadora deben incluir todos los aspectos, desde la especificación hasta la implementación. Trabajos recientes, en particular aquellos en torno al protocolo TLS 1.3, han demostrado que este enfoque es ahora factible. Sin embargo, esto sigue requiriendo un cuidadoso diseño conjunto de prueba y código que sólo puede ser realizado por expertos. Aprovechar estas pruebas para obtener un código más general y propiedades de seguridad más complejas, por ejemplo, propiedades que garanticen la privacidad del usuario, sigue siendo un gran reto.

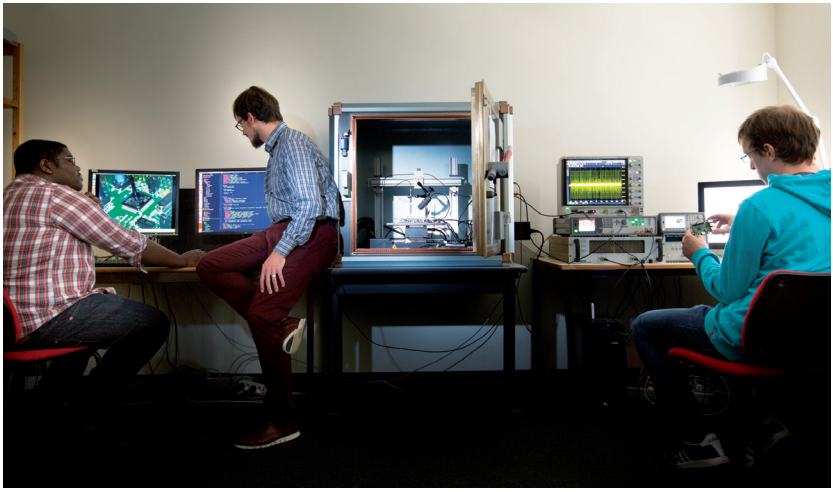
→ *Seguridad para IoT.* La seguridad para la IoT es un reto importante. Los ataques siguen siendo relativamente fáciles (muchos dispositivos no se han diseñado teniendo en cuenta la seguridad), invasivos (por ejemplo, en nuestras vidas), y tienen un gran impacto potencial debido al gran número de dispositivos disponibles, lo que aumenta la superficie de ataque y facilita los ataques de denegación de servicio descentralizado (DDoS). Las líneas de investigación son numerosas: mencionemos, por ejemplo, el deseo de actualizar de forma segura el software embebido en un dispositivo, la necesidad de primitivas criptográficas ligeras adaptadas a recursos limitados, el análisis de la seguridad de las nuevas tecnologías inalámbricas de área amplia y de baja potencia, la detección y mitigación de intrusiones o de dispositivos con funcionamiento anómalo, o la necesidad de marcos y protocolos de diseño seguro para facilitar el desarrollo de los dispositivos IoT.

→ *Proteger la privacidad de los ciudadanos.* Nuestro mundo conectado experimenta un crecimiento sin precedentes en lo que respecta a la recopilación de datos personales, cada vez más intrusivos, ya sea mientras se navega por la

web, se utiliza un smartphone o se conduce un automóvil conectado. La falta de transparencia, el hecho de que muchos servicios y dispositivos se comporten como cajas negras, y la falta de control del usuario son problemas importantes. ¿Cómo se puede expresar el consentimiento o la oposición si no existe información ni interfaz de usuario? La identificación de estos comportamientos ocultos, que requiere análisis de flujo de datos, se ve dificultada por el número, la complejidad y la diversidad de las aplicaciones y tecnologías de comunicación subyacentes. Se necesitan actividades de investigación transversales ambiciosas e innovadoras para aportar transparencia, poner de relieve las buenas y malas prácticas y permitir a los reguladores hacer cumplir la legislación sobre protección de datos.

Es esencial incluir la seguridad desde el principio en el diseño de sistemas. Lo mismo ocurre con la ciber-resiliencia: los ciberataques masivos son una amenaza cada vez mayor y el principio de seguridad en el diseño debe aplicarse también a la resiliencia de las redes e infraestructuras digitales críticas.

La ciberseguridad carece a menudo de incentivos, ya que sus beneficios son difíciles de comprender. Esto se debe a menudo a la falta de experiencia. Efectivamente, los académicos, que tienen una gran experiencia en la mayoría de los ámbitos de la ciberseguridad, suelen estar poco representados en los comités asesores nacionales o industriales. La educación es, por tanto, esencial para la seguridad y hay que hacer un gran esfuerzo de difusión para todos los públicos: desde los profesores, los investigadores, los agentes industriales y especialistas, hasta los ciudadanos comunes, incluidos los niños.



Laboratorio de alta seguridad (LHS) en el Centro de Investigación Inria de Rennes – Bretagne Atlantique © *Inria / Photo C. Morel*

Acerca de este libro blanco

Nuestros objetivos

El objetivo principal de este libro blanco de Inria sobre ciberseguridad es el de detallar la visión de Inria sobre los retos de la ciberseguridad. Para ello, incluimos una visión general de los temas de investigación académica en ciberseguridad y, en particular, una cartografía de la investigación existente sobre ciberseguridad en Inria. Asimismo, aprovechamos la oportunidad para formular recomendaciones generales en el ámbito de la ciberseguridad. Hemos optado por llevar a cabo todos estos sub-objetivos de forma paralela en un único documento unificado, sin una distinción constante entre los objetivos. Por lo tanto, aunque pretendemos una cobertura completa de la investigación en ciberseguridad, de manera intencional el nivel de detalle no es uniforme: se enfatizan los dominios en los que Inria tiene una posición fuerte o, en el extremo opuesto, en los que la contribución de Inria debería incrementarse; los campos menos estructurados también reciben una presentación más completa.

Este libro blanco está escrito con la intención de llegar a un público lo más diverso posible y permitir distintos niveles de lectura. Incluye presentaciones técnicas de los distintos ámbitos de la ciberseguridad y una descripción detallada del trabajo realizado en los equipos de Inria que podría ser de interés para los expertos en ciberseguridad o a quienes busquen información detallada sobre un subdominio concreto. Adicionalmente, incluye información más accesible tipografiada en cuadros de texto: las secciones que cubren el material técnico incluyen un **[Resumen]**, ejecutivo, así como una lista de **[Equipos Inria]** que trabajan en cada área con un breve resumen de sus actividades relacionadas. Otros elementos de información adicional son recuadros con **[Secciones Destacadas]**, **[Notas]** pedagógicas y **[Desafíos de Investigación]**

Por lo tanto, el lector no experto puede centrarse primero en los cuadros de texto, y sólo profundizar en el texto completo cuando lo considere necesario. Los retos se describen *in situ*, pero también se recogen en un capítulo específico (apartado 8.1) para mayor comodidad.

La metodología

Este libro blanco es un trabajo colaborativo, con contribuciones de muchas personas de Inria y sus socios. La redacción de este libro fue coordinada por un grupo de trabajo formado por Steve Kremer, Ludovic Mé, Didier Rémy y Vincent Roca. Dada la amplitud del tema, contaron con la ayuda de muchos otros investigadores que aportaron al grupo de trabajo sus conocimientos sobre el tema.

En un segundo paso, los capítulos han sido revisados por investigadores que trabajan en las respectivas áreas. Las citas de artículos científicos están intencionadamente restringidas a los trabajos más importantes, en lugar de intentar ofrecer una bibliografía extensa. Este libro blanco se ha realizado bajo la supervisión de la *"Cellule de veille et prospective"* –la unidad de vigilancia científica y prospectiva de Inria– y forma parte de una serie de libros blancos.

Esquema

El resto de este libro blanco ofrece una visión general de las áreas de investigación en ciberseguridad y, en particular, de las actividades de los equipos de Inria. Indudablemente, existen muchas maneras de presentar las actividades en ciberseguridad, clasificándolas por metodologías, sub- comunidades, dominios de aplicación, etc. En este libro seguimos un camino sinuoso, pues nos pareció la mejor manera para presentar las actividades de Inria en materia de ciberseguridad.

He aquí un resumen de nuestra trayectoria en materia de ciberseguridad:

Capítulo 1: INTRODUCCIÓN

Definimos el contexto: Comenzamos con el alcance de la ciberseguridad, se discuten los problemas y los retos, y se dan ejemplos de ataques y sus consecuencias. Adicionalmente, se analizan las propiedades clave de la seguridad, así como algunas consideraciones legales (por ejemplo, la regulación de la ciberseguridad) y de soberanía.

Capítulo 2: AMENAZAS

Los trabajos sobre ciberseguridad suelen empezar por definir el "modelo de ataque": es decir, las capacidades de un atacante. En este capítulo analizamos diferentes amenazas que pueden dirigirse al hardware, la red, el sistema operativo, las aplicaciones o incluso los propios usuarios. Adicionalmente, repasamos algunos proyectos de investigación cuyo objetivo es una mejor comprensión de estas amenazas.

Capítulo 3: CRIPTOGRAFÍA

La criptografía desempeña un papel esencial y constituye la base de la ciberseguridad. En este capítulo cubrimos todos los aspectos de la criptografía, desde el diseño de las primitivas básicas hasta los protocolos más complejos que proporcionan garantías de alto nivel en cuanto a la seguridad de las comunicaciones y las transacciones.

1. https://www.inria.fr/institut/strategie

Capítulo 4: SERVICIOS DE SEGURIDAD

Necesitan servicios de seguridad adicionales para diseñar sistemas operativos, aunque se construyan a menudo sobre primitivas y protocolos criptográficos. Estos servicios son presentados en este capítulo, en el que abordamos mecanismos de seguridad que pueden prevenir o mitigar las amenazas y los ataques a los sistemas de información y sus componentes, incluidos el hardware, las redes y los sistemas operativos.

Capítulo 5: PRIVACIDAD

Hoy en día, la privacidad se considera una parte intrínseca de la ciberseguridad. Aunque la privacidad se basa en primitivas y protocolos criptográficos, también tiene sus propias propiedades, técnicas y metodología. En este capítulo nos centramos en la privacidad, abarcando tanto los aspectos técnicos como los legales, económicos y sociológicos.

Capítulo 6: APLICACIONES SENSIBLES EN SEGURIDAD

Mientras que los capítulos anteriores se centran en servicios y herramientas específicas, aquí adoptamos el enfoque opuesto, analizando un conjunto seleccionado de aplicaciones sensibles en seguridad y analizando las cuestiones de seguridad específicas que plantean.

Capítulo 7: LA CIBERSEGURIDAD EN FRANCIA

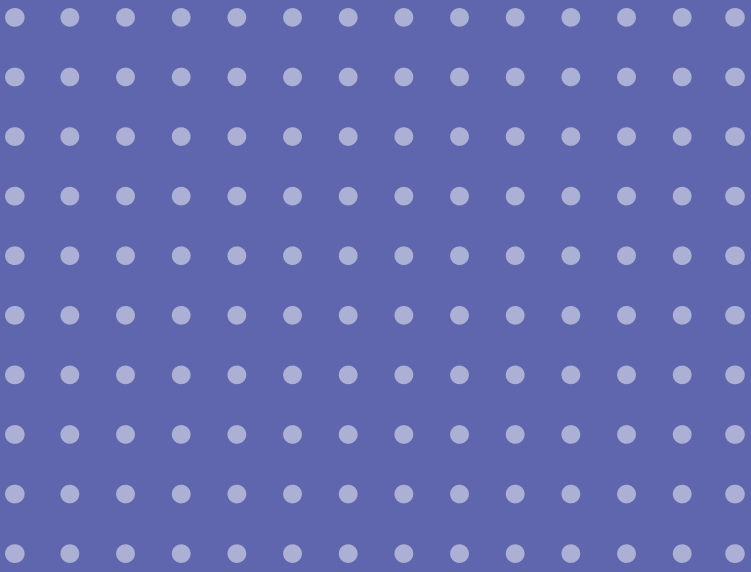
Por último, en el capítulo 7, ofrecemos una visión general de las actividades de ciberseguridad en Inria y su posicionamiento en Francia.

Finalmente, la conclusión del libro contiene recomendaciones generales para llevar a cabo.

Índice

1

Introducción



© Inria / ilustración Clod

1.1 La ciberseguridad, una preocupación fundamental

Una grave preocupación para nuestras sociedades

Mientras que la palabra *ciberseguridad* apenas era conocida por el público en general hace dos décadas, se ha convertido en un tema recurrente en los medios de comunicación públicos recalcado por los frecuentes ataques a la ciberseguridad, el descubrimiento de nuevas brechas o el descubrimiento de la existencia de una vigilancia masiva por parte de grandes empresas y organismos estatales.

El mundo ha cambiado, y lo ha hecho rápidamente. La ciberseguridad se ha convertido en una preocupación general para todos: ciudadanos, profesionales, políticos y, en general, todos los responsables de la toma de decisiones. Asimismo, se ha convertido en una gran preocupación para nuestras sociedades, que deben protegerse contra los ataques de ciberseguridad con medidas tanto preventivas como reactivas, lo que implica ejercer mucha vigilancia, a la vez que se debe preservar nuestra libertad y evitar la vigilancia masiva.

Los ciberataques pueden ser llevados a cabo por delincuentes, pero también por los Estados con el propósito de realizar espionaje industrial, de causar daños económicos para ejercer presión, o de infligir daños reales a infraestructuras como un acto de guerra.

Los Estados y sus infraestructuras críticas interconectadas son vulnerables. Los ciberataques también ponen a las empresas –de todos los tamaños– en alto

riesgo. Los daños económicos causados por ciberataques exitosos pueden ser considerables. Sin embargo, nuestro nivel de protección todavía se considera muy insuficiente en comparación con los riesgos y los daños potenciales. Aunque nuestra concienciación está mejorando y las medidas de protección están aumentando, todavía lo hacen a un ritmo demasiado lento. Esto se debe, en parte, a la falta de incentivos: la ciberseguridad es una inversión cuyos beneficios son, por lo general, difíciles de comprender, ya que solo resulta rentable cuando un ataque que podría haber tenido éxito fracasa, y esto es difícil de medir. Esta lentitud se debe también a la falta de conocimientos técnicos a todos los niveles.

La ciberseguridad es, por tanto, una cuestión económica y de soberanía para muchos Estados, incluida Francia.

Un reto para la era digital

La digitalización de nuestra sociedad está cambiando por completo el uso de los sistemas informáticos. Hasta principios de la década de los 90, los sistemas informáticos estaban poco conectados. Pocas personas tenían computadoras personales en casa, y éstas rara vez estaban conectadas a Internet. El spam por correo electrónico apenas existía. La ciberseguridad era principalmente una preocupación para los Estados y las grandes empresas, incluida la industria financiera.

Desde finales de la década de los 90, la situación ha cambiado por completo. El crecimiento de la web y la aparición del ADSL y los smartphones llevaron rápidamente conexiones de Internet baratas y rápidas a casi todos los hogares de los países desarrollados y revolucionaron el papel de Internet. Los teléfonos inteligentes y la tecnología 3G/4G también han extendido el acceso a Internet a los países en desarrollo. Una enorme proporción de la población está ahora continuamente conectada a Internet, utilizando una cantidad asombrosa de servicios diferentes. Al mismo tiempo, nos hemos visto permanentemente expuestos a ataques, con nuestros datos sensibles en riesgo de ser robados o dañados. Adicionalmente, vivimos con el riesgo de filtrar por error y de forma irreversible nuestra información privada en Internet. La ciberseguridad es hoy en día un problema muy serio para todos, desde los ciudadanos, pequeñas y grandes empresas a organismos administrativos y estatales. Por ejemplo, un ataque exitoso de ransomware (secuestro de datos) a una pequeña empresa que no tiene copias de seguridad (bien aisladas) puede poner en peligro el negocio de la empresa, o incluso llevarla a la quiebra. La empresa puede perder el acceso a todos sus pedidos, y los listados de sus proveedores y clientes; la empresa en cuestión quizás se vea obligada a pagar una enorme cantidad de dinero por la clave de descifrado, que en el peor de los casos puede que no llegue a entregarse.

La superficie de ataque ha aumentado considerablemente debido al número de dispositivos, algunos de ellos con poca seguridad o completamente inseguros, y la mayoría de ellos manejados por usuarios desinformados. Esto ha aumentado

enormemente las posibilidades de que un ataque tenga éxito, en beneficio del atacante, haciendo que la ciberdelincuencia sea más provechosa. Además, el ataque puede lanzarse desde cualquier parte del mundo.

Es probable que la situación se agrave en los próximos años con la difusión de los dispositivos conectados que forman la Internet de las Cosas (IoT), la cual está todavía en su etapa inicial. El número de dispositivos digitales conectados puede aumentar al menos en otro orden de magnitud. Esta interconexión expone a todo el mundo al nivel más bajo de seguridad de las máquinas de confianza a las que cualquiera está conectado, ya que ellos mismos pueden servir como ataque. Al confiar en los bajos niveles de protección, no solo aumentamos el riesgo para nosotros mismos, sino también el riesgo de contribuir a la propagación de un ataque a gran escala. Esto significa que cada actor tiene una responsabilidad y una obligación moral hacia la comunidad de implementar medidas de protección suficientes. La IoT aumenta tanto la superficie de ataque como el aspecto de contaminación de las violaciones de la ciberseguridad, lo que exige niveles de seguridad aún más elevados.

Difusión a través de la informática

Dado que un sistema completo es tan fuerte y seguro como su eslabón más débil, la interconexión e interacción de máquinas y aplicaciones (que a menudo se ejecutan en un entorno distribuido) hace que la ciberseguridad sea una cuestión central para casi todo el software actual. Por tanto, no es sorprendente observar la difusión de conocimientos en materia de ciberseguridad a otros ámbitos de la informática.

Existen varias razones que convergen para explicar esta evolución, que conduce a la fertilización cruzada interdisciplinaria. Por ejemplo, la ciberseguridad ha sido un ámbito de especial interés para la investigación de los métodos formales: la ciberseguridad necesitaba urgentemente conocimientos especializados en el ámbito de los métodos formales y, al mismo tiempo, constituía un territorio extremadamente desafiante y estimulante para los investigadores en métodos formales. En la actualidad, los métodos formales están bien implantados en el ámbito de la ciberseguridad, especialmente en Francia, aunque se trata de un campo de investigación aún en crecimiento (véanse los apartados 3.3.2 y 3.3.3).

En el ámbito de las bases de datos, los investigadores están cada vez más expuestos a los problemas de seguridad. Al principio los investigadores eran usuarios de la ciberseguridad, importando técnicas estándar a sus ámbitos de investigación, pero rápidamente se han convertido en colaboradores activos de la ciberseguridad. Por ejemplo, la comunidad de las bases de datos ha propuesto nuevos paradigmas, como la nube privada, para aplicar la privacidad por diseño.

En otras áreas, como la seguridad del sistema, la computación distribuida y los servicios de red, las razones son más complejas; pero en cada caso, esta evolución

refleja la creciente exposición de estos dominios a los problemas de seguridad. Más recientemente, se han producido nuevas y fructíferas interacciones entre la ciberseguridad y el aprendizaje automático, en ambas direcciones: por un lado, los métodos de aprendizaje automático se están aplicando a la seguridad, especialmente en la seguridad reactiva (apartados 4.4 y 4.6). Por otro lado, el aprendizaje automático plantea nuevos problemas de seguridad y privacidad (véase el apartado 6.3.3).

Difusión entre otras ciencias

Y lo que es más importante, la ciberseguridad se está convirtiendo en una preocupación importante para ámbitos de aplicación ajenos a la informática que utilizan dispositivos informáticos o servicios digitales de manera crítica: la cibersalud, la medicina (apartado 6.3.1), la robótica (apartado 6.3.2), las centrales eléctricas y de suministro de agua, las ciudades inteligentes y, en general, todas las infraestructuras esenciales (apartado 6.2.2).

Esta interdependencia entre los subdominios de las ciencias de la computación en lo que respecta a las cuestiones de seguridad es a la vez un verdadero reto y una oportunidad que Inria debe aprovechar, ya que está presente en la mayoría de estos subdominios. La difusión de la ciberseguridad en otras ciencias es una oportunidad para aplicar la investigación en ciberseguridad, para la que habrá una permanente demanda, y expectativas.

1.2 El alcance de la ciberseguridad

Wikipedia define la ciberseguridad de la siguiente manera:

La seguridad informática, también conocida como ciberseguridad o seguridad TI, es la protección de los sistemas informáticos contra el robo o daño de su hardware, software o información, así como contra la interrupción o el desvío de los servicios que prestan.

Sin embargo, la noción exacta de ciberseguridad varía según el contexto[2]. La seguridad en general incluye tanto la ciberseguridad como la seguridad física. No obstante, la ciberseguridad requiere alguna forma de seguridad física, ya que el acceso físico a los sistemas informáticos permite toda una clase de ataques. A la inversa, la seguridad física puede depender de la ciberseguridad en la medida en que utilice sistemas informáticos, por ejemplo, para controlar algún espacio físico o mantener una base de datos de personas autorizadas. En cualquier caso, la diferencia entre la ciberseguridad y la seguridad física tiene que estar siempre clara, y

2. ENISA, "Denition of Cybersecurity: Gaps and overlaps in standardization", version 1.0, December 2015. https://www.enisa.europa.eu/publications/definition-of-cybersecurity.

en lo sucesivo sólo abordaremos el tema de la ciberseguridad. Efectivamente, en muchos apartados nos limitaremos a utilizar la palabra seguridad para referirnos a la ciberseguridad.

[Nota] Seguridad física frente a ciberseguridad

La seguridad física y la ciberseguridad son de naturaleza bastante diferente.

La información digital es inmaterial: duplicar e intercambiar datos y códigos con cualquier persona en cualquier parte del mundo es hoy un proceso trivial, extremadamente rápido y con un coste casi nulo. De ahí que un ataque o un programa malicioso lanzado por una sola persona pueda extenderse por todo el mundo, a gran escala, en menos de una hora.

La información digital es de naturaleza discreta: un solo bit flip puede introducir un fallo crítico y convertir un sistema que funciona perfectamente en uno que no funciona, el cual es entonces más vulnerable a ser atacado. Esto contrasta con las leyes de la física, que tienden a ser continuas a nivel macroscópico y suelen permitir observar una lenta deformación de una estructura antes de que alcance su punto de ruptura. La información digital ignora las fronteras, e incluso puede aprovecharse de las contradicciones entre las legislaciones de diferentes países o su inadaptación a la era digital.

Esto hace que la ciberseguridad sea mucho más difícil de ejercer que otras formas de seguridad.

Prevención frente a seguridad

La prevención (o safety) del software se refiere a la ausencia de comportamientos erróneos, tanto en situaciones normales como excepcionales, pero también en un entorno neutral cuando nadie intenta atacar intencionadamente el sistema. La prevención en el contexto del software no es sólo una cuestión de detección de errores *(o bugs)*: también requiere un análisis de las posibles fuentes de mal comportamiento y de cómo manejarlas de forma segura. Esto requiere una especificación del comportamiento esperado del software, incluyendo un modelo de entorno, y alguna justificación de cómo o por qué el software respeta su especificación.

En cambio, la seguridad (*o security*) del software tiene como objetivo la ausencia de comportamientos erróneos en un entorno adverso, en el que un atacante intenta de manera intencional hacer un mal uso de un sistema, poniéndolo en un estado erróneo que no forma parte de su especificación prevista. La seguridad también puede abordarse modelando el entorno, pero esto es mucho más difícil de conseguir de forma exhaustiva, porque los atacantes no cumplen con las reglas predefinidas, sino que buscan continuamente medios de ataque previamente desconocidos. Por lo tanto, la seguridad también requiere que nos mantengamos

al día sobre los avances de los atacantes en todos los ámbitos (fallos o violaciones del software, algoritmos y técnicas, capacidades del hardware, etc.). Un enfoque complementario consiste en describir flujos de ejecución normales y supervisar la ejecución, para dar la alarma y reaccionar adecuadamente cuando alguna trayectoria se salga de las ejecuciones normales.

[Nota] Prevención frente a seguridad

Los términos seguridad y prevención se utilizan a veces de forma errónea. La prevención se refiere a las amenazas accidentales, debidas a comportamientos internos erróneos o a un mal uso no intencionado del sistema, mientras que la seguridad se refiere a las amenazas intencionales. La prevención se refiere a la tolerancia a fallos, mientras que la seguridad se refiere a la resistencia a los ataques. Por ejemplo, un coche puede estrellarse debido a una especificación del software o a un error de implementación (problemas de prevención), o porque un atacante tome el control remoto del vehículo (un problema de seguridad).

A pesar de estas diferencias, la prevención y la seguridad suelen ir unidas. Dado que la seguridad funciona en modo adversario, también debe tener en cuenta las amenazas accidentales que pueden ser explotadas por el atacante. Por lo tanto, la seguridad es un requisito más importante que la prevención. Sin embargo, en muchas situaciones, el software está expuesto al mundo exterior, normalmente conectado a Internet, donde los ataques son la norma y la prevención sin seguridad no tendría mucho sentido.

La prevención y la seguridad también comparten mucho en sus metodologías. El tratamiento de la prevención de los grandes sistemas de software que interactúan con el mundo físico, como los sistemas ciberfísicos (también conocidos como CPS, véase el apartado 6.2), ha dado lugar a algunas metodologías bien establecidas. Hay que empezar con una fase de análisis de riesgos de prevención, en la que se exploran todas las situaciones que pueden tener consecuencias catastróficas. Se pueden utilizar representaciones como los árboles de fallas para describir sistemáticamente estas situaciones.

A continuación, los riesgos identificados se cuantifican para estimar la probabilidad de que se produzcan estas situaciones. Garantizar la prevención significa asegurar que esta probabilidad se mantiene por debajo de un umbral determinado. Por supuesto, un primer paso para satisfacer la propiedad de prevención es garantizar la ausencia de fallos internos ***(bugs)*** en el software, ya que estos son la causa principal de los fracasos. En primer lugar, se escribe una especificación del software para describir el comportamiento esperado del mismo y luego se demuestra que la implementación satisface la especificación. Por desgracia, es posible que no se hayan tenido en cuenta todos los casos en la especificación. Además, puede haber fallos externos (por ejemplo, un valor erróneo procedente

de un sensor externo) que no se hayan tenido en cuenta en la especificación del software, y que pueden provocar desastres. De ahí que debamos utilizar también mecanismos adicionales, basados esencialmente en detección dinámica de errores y mecanismos de recuperación, utilizados para tratar los errores causados por fallos externos antes de que provoquen consecuencias catastróficas.

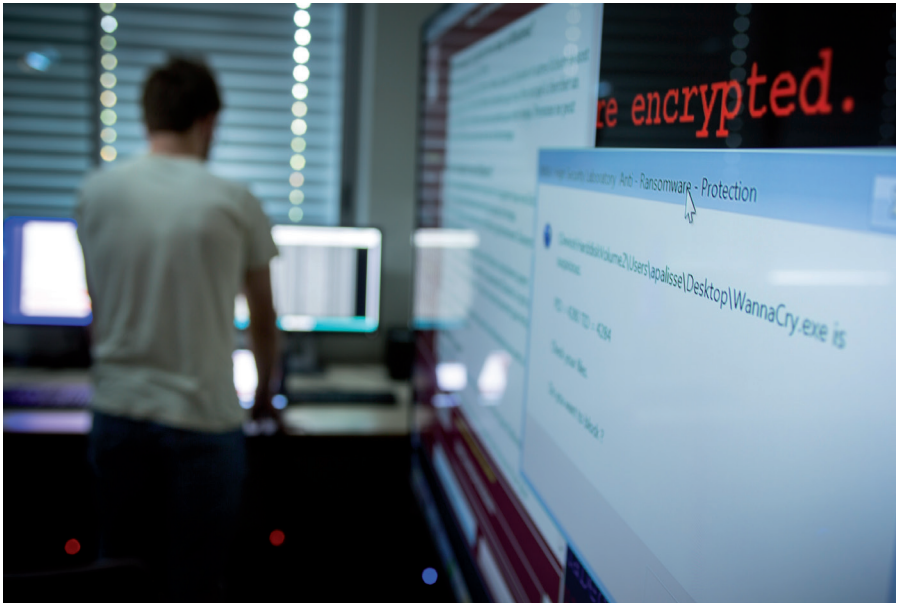
Un enfoque similar se aplica a la seguridad. Un análisis de riesgos de seguridad (véase el apartado 2.4) sustituye al análisis de riesgos de prevención. Aunque no es posible razonar estadísticamente para construir un sistema inatacable en el caso de la seguridad, sigue siendo útil asegurarse de que no hay fallos (al menos de algún tipo) en el software, por ejemplo mediante el mismo enfoque formal que es seguido para la prevención, porque los ataques suelen basarse en las vulnerabilidades que se derivan de un fallo subsistente. La supervisión adopta la forma de detección de ataques dinámicos y mecanismos de recuperación. Esto implica un modelo del atacante, que debería cubrir al menos todos los tipos de ataques conocidos, por ejemplo en forma de un ataque basado en firmas. Desde este punto de vista, el principio de prevención por diseño se convierte en el principio de seguridad por diseño, lo que significa que el software debe diseñarse desde los cimientos para ser seguro[3]. Esto ha dado lugar a principios de diseño como las recomendaciones OWASP[4].

Sin embargo, la seguridad y la prevención siguen siendo ámbitos distintos y diferentes, construidos sobre hipótesis distintas, y los mecanismos de protección contra las amenazas accidentales e intencionales suelen ser complementarios. En este libro blanco limitamos nuestra atención a la seguridad.

1.3 Algunos ejemplos y lecciones aprendidas

Desgraciadamente, los incidentes de ciberseguridad son frecuentes y a menudo aparecen en los titulares. Aquí describimos algunos ejemplos ilustrativos, con el fin de destacar la enorme diversidad de ataques. Algunos de ellos tienen como objetivo entidades bien identificadas, como TV5 Monde y la empresa Dyn, aunque los ataques utilizaron técnicas totalmente diferentes. En el otro extremo del espectro, el ransomware WannaCry fue dirigido a una gran cantidad de personas, propagándose de manera viral. Se sabe que algunos sistemas de votación electrónica son vulnerables y, en varias ocasiones, los investigadores de seguridad han puesto de manifiesto su insuficiencia mediante ataques de prueba de concepto. El siguiente ejemplo ilustra cómo las bases de datos anónimas pueden a veces ser atacadas, desvelando identidades físicas. Los dos últimos ejemplos destacan, en el primer caso, dos ataques de software dirigidos al hardware que explotan técnicas avanzadas de optimización del rendimiento del procesador, y

3. Ver Wikipedia https://en.wikipedia.org/wiki/Secure_by_design
4. https://www.owasp.org/index.php/Security_by_Design_Principles



Investigación sobre el malware – © Inria / Foto C. Morel

en el segundo, las debilidades de algunos dispositivos de la Internet de las Cosas y su explotación.

Ataque dirigido a TV5 Monde: el 9 de abril de 2015, la cadena de televisión francesa TV5 Monde fue víctima de un gran sabotaje. Hacia las 21:00 horas, el sitio web y los canales de las redes sociales (Facebook, Twitter, YouTube) fueron alterados. Una hora más tarde, la infraestructura de red dejó de estar operativa y se interrumpió la emisión, lo que provocó un apagón televisivo completo, lo peor que le puede pasar a una cadena de televisión. La Agencia Nacional Francesa de Ciberseguridad (ANSSI) descubrió posteriormente que el ataque había sido cuidadosamente planificado[5]. Los atacantes se conectaron por primera vez en enero, utilizando un nombre de usuario y una contraseña robados. Esto les permitió acceder a la red interna, recopilar documentos internos que contenían información sobre la infraestructura de red y las cuentas existentes y explotar servicios no configurados que todavía dependían de las cuentas y contraseñas por defecto. El borrado de firmware en la infraestructura de red (routers y switches) provocó entonces la caída, haciendo imposible un simple reinicio.

Ataques de denegación de servicio desde la red de bots Mirai de dispositivos domésticos: el objetivo del malware Mirai es convertir los dispositivos domésticos

5. https://static.sstic.org/videos2017/SSTIC_2017-06-09_P09.mp4

vulnerables (como cámaras IP, impresoras, monitores de bebés o routers domésticos) en bots controlados de forma remota que posteriormente pueden utilizarse para lanzar ataques de denegación de servicio a gran escala. Esto es lo que ocurrió el 21 de octubre de 2016, cuando esta red de bots atacó los servidores de nombres de la empresa Dyn (los que son utilizados, por ejemplo, para traducir los nombres de dominio a direcciones IP, véase el apartado 2.2). Este ataque provocó el bloqueo de muchos sitios web en todo el mundo durante varias horas.

El ransonware WannaCry: el viernes 12 de mayo de 2017, el ransomware WannaCry se propagó por todo el mundo, infectando más de 230.000 computadoras en más de 150 países en un solo día (fuente Wikipedia). Este ransomware se dirige a computadoras con el sistema operativo Microsoft Windows, con importantes consecuencias para sus propietarios: tras infectar una computadora, el ransomware cifra los datos y muestra una nota para informar al usuario, donde pide un pago en bitcoin a cambio de la clave de descifrado. Este ransomware se considera un gusano, ya que escanea en busca de sistemas vulnerables y luego se replica en estos nuevos objetivos.

Vulnerabilidades del voto electrónico: en los últimos años, varios países europeos (Estonia, Francia, Noruega y Suiza) celebraron elecciones políticas legalmente vinculantes que permitieron a una parte de los votantes emitir su voto a distancia a través de Internet. Los franceses que viven en el extranjero pudieron votar por Internet en las elecciones parlamentarias de junio de 2012. Un ingeniero demostró que era posible escribir un software malicioso que podía cambiar el voto emitido, sin que el votante pudiera saberlo. En las elecciones parlamentarias de Estonia de 2011, un informático llamado Pihelgas denunció un ataque similar y llevó a cabo un experimento en la vida real con sujetos plenamente conscientes.

Reidentificación en la base de datos anonimizada de consultas de búsqueda web de AOL: como informó el New York Times[6], AOL publicó una base de datos anonimizada que contenía más de 20 millones de consultas de búsqueda web. Aunque los datos estaban anonimizados, los usuarios pudieron ser identificados después de algunas investigaciones, revelando así todas sus consultas de búsqueda personales. En términos más generales, la anonimización de bases de datos es una tarea compleja que presenta escollos y que requiere encontrar un equilibrio adecuado entre utilidad y privacidad.

Las vulnerabilidades Spectre y Meltdown: el 3 de enero de 2018 se publicaron simultáneamente dos vulnerabilidades de hardware, Spectre y Meltdown. Ambas vulnerabilidades explotan la ejecución especulativa (y en particular la predicción de saltos), una técnica de optimización en los procesadores modernos. Para evitar los ciclos ociosos del procesador, por ejemplo, mientras se espera el resultado de un acceso a la memoria, los procesadores pueden realizar una ejecución fuera de

6. New York Times, "A Face Is Exposed for AOL Searcher No. 4417749", August 9, 2006 http://www.nytimes.com/2006/08/09/technology/09aol.html

orden. Un salto puede entonces ser ejecutado especulativamente, mientras espera la evaluación de una condicional. Si el salto se ejecuta erróneamente, los resultados se descartan. Sin embargo, aunque se descarten los resultados, un acceso a la memoria deja un rastro en la caché. La idea de los ataques Spectre y Meltdown es forzar un acceso prohibido a la memoria. Normalmente, los desbordamientos de búfer se evitan mediante comprobaciones del tamaño del búfer. Sin embargo, estas comprobaciones pueden ser evitadas haciendo que la predicción de salto prediga erróneamente la prueba. Entonces, se puede utilizar un ataque a la caché para comprobar qué zona de la memoria se ha ejecutado. (Estos ataques simplemente miden el tiempo necesario para acceder a una dirección de memoria concreta). Los ataques son especialmente graves porque se aprovechan del diseño de los procesadores modernos y no se pueden parchear simplemente con una actualización del software. Además, la ejecución especulativa está en el núcleo del diseño de los procesadores modernos, y es poco probable que los fabricantes de procesadores dejen de utilizarla.

→ ***Luces inteligentes que provocan ataques de epilepsia:*** investigadores del Instituto Weizmann de Ciencias han demostrado que es posible hackear las luces inteligentes más comunes y hacerlas parpadear a una frecuencia que puede desencadenar convulsiones epilépticas[7].[ER16] Este ataque es interesante porque al convertir objetos tradicionalmente desconectados (en este caso, las bombillas) en objetos inteligentes, se pueden utilizar de forma indebida para crear un ataque inesperado. Este ataque en particular aprovecha una combinación de varios fallos. En primer lugar, al inicializar el controlador de luces inteligente, la contraseña que permite al controlador conectarse a la wifi local se envía sin cifrar y es fácilmente detectable. En segundo lugar, las luces aceptan comandos de cualquier dispositivo en la wifi local sin un mecanismo de autenticación adecuado. En tercer lugar, el controlador no verifica la longitud de los comandos que recibe, lo que permite la concatenación de múltiples comandos, eludiendo el límite de comandos que se pueden enviar por segundo. Por último, el ataque se basa en opciones de API no documentadas, lo que permite a los atacantes crear un efecto estroboscópico.

Lecciones aprendidas

Estos ejemplos ponen de manifiesto varios aspectos clave de la seguridad:

→ ***La seguridad es una piedra angular esencial en un mundo digital que impregna cada vez más todos los aspectos de nuestra vida cotidiana, pública y privada.*** Sin seguridad, el mundo se derrumba. Ataques como WannaCry han

7[ER16]. E. Ronen, A. Shamir. Extended functionality attacks on iot devices: The case of smart lights. In *IEEE European Symposium on Security and Privacy (EuroS&P'16)*, 2016.

impactado profundamente a ciudadanos, empresas privadas y organizaciones no preparadas, amenazando sus actividades.

→ ***Todos los ámbitos de nuestro mundo digital están afectados***, incluidos los dispositivos embebidos que están omnipresentes en nuestros hogares "inteligentes" y en los controladores de producción industrial (incluidos los de infraestructuras críticas como el suministro de energía y agua). Dado que todos ellos están conectados a Internet, la seguridad es una grave preocupación, como demuestra el ataque a las luces inteligentes mencionado anteriormente. El ejemplo de la red de bots Mirai pone de manifiesto que todos los dispositivos electrónicos necesitan ser seguros. Aunque esto es bien sabido en el caso de las computadoras, dista mucho de ser obvio en el caso de otros objetos, en particular los dispositivos embebidos que forman la Internet de las Cosas (IoT), ya sea porque son autónomos, tienen una batería pequeña, una potencia de procesamiento limitada o están mal conectados. Además, la incapacidad de los dispositivos de la IoT para aplicar actualizaciones y parches de software es un auténtico problema. Por otra parte, las actualizaciones de software pueden ser en sí mismas objeto de ataques. Todos estos aspectos siguen siendo objeto de investigación activa.

→ ***La educación es esencial para la seguridad***. El ataque de WannaCry se basó en un exploit del sistema operativo que había sido corregido en una actualización de Windows dos meses antes. Esto solo afectó a usuarios finales y administradores de sistemas no preparados que no actualizaron sus computadoras a tiempo, sin darse cuenta de su importancia. La seguridad suele considerarse compleja, lo que limita mecánicamente su uso. La seguridad usable, destinada a facilitar el uso de la seguridad por parte de los usuarios finales, es un ámbito de investigación importante y activo que está estrechamente relacionado con la educación y la concienciación en materia de seguridad.

→ ***La seguridad de un sistema siempre está limitada por la de su componente más débil.*** Aunque los componentes centrales de seguridad (por ejemplo, las primitivas criptográficas) rara vez son atacados, no puede decirse lo mismo de las implementaciones de software de los protocolos y servicios criptográficos. En el caso de WannaCry, el ataque se basó en un exploit del protocolo SMB de Windows (el primer eslabón débil), que fue suficiente para tomar el control total de la computadora, sin importar qué otras protecciones del sistema operativo estuvieran en uso. El segundo eslabón débil fueron los usuarios, que al no actualizar sus computadoras hicieron que el ataque tuviera éxito. El ataque a TV5 Monde fue posible, en primer lugar, gracias a la ingeniería social y, posteriormente, al uso de contraseñas y nombres de usuario no modificados en varios equipos técnicos.

→ ***La oscuridad no aumenta la seguridad.*** A veces, la gente cree que ocultar las partes internas de un sistema o mecanismo de seguridad aumentará la seguridad. Sin embargo, ahora sabemos que, por el contrario, los principios de diseño abierto mejoran la seguridad de un sistema. En criptografía, este hecho se conoce como el principio de Kerckhoffs, y se remonta al siglo XIX: Un criptosistema debe ser seguro aunque todo lo relacionado con el sistema, excepto la clave, sea de dominio público. Este principio debería aplicarse también a otros sistemas. Un diseño abierto y un sistema bien documentado facilitarán las revisiones de seguridad por parte de los expertos. Los atacantes suelen ser capaces de aplicar ingeniería inversa a los sistemas, y la "seguridad por medio de la oscuridad" sólo da una falsa sensación de seguridad. Por ejemplo, el ataque a las luces inteligentes explotó una funcionalidad no documentada.

→ ***Los sistemas grandes y complejos no se pueden validar totalmente mediante la inspección humana: se necesitan herramientas de verificación automática para encontrar fallos en los protocolos de seguridad, así como en la implementación.*** El componente SMB al que apuntaba su ataque WannaCry tiene una larga historia detrás. A pesar de ello, se pudo identificar un fallo de seguridad que hizo posible el ataque. La creciente complejidad de cada componente individual, y la compleja composición de componentes en grandes sistemas interdependientes, requieren herramientas de validación de seguridad avanzadas y automáticas, lo que tradicionalmente constituye un tema de investigación muy activo.

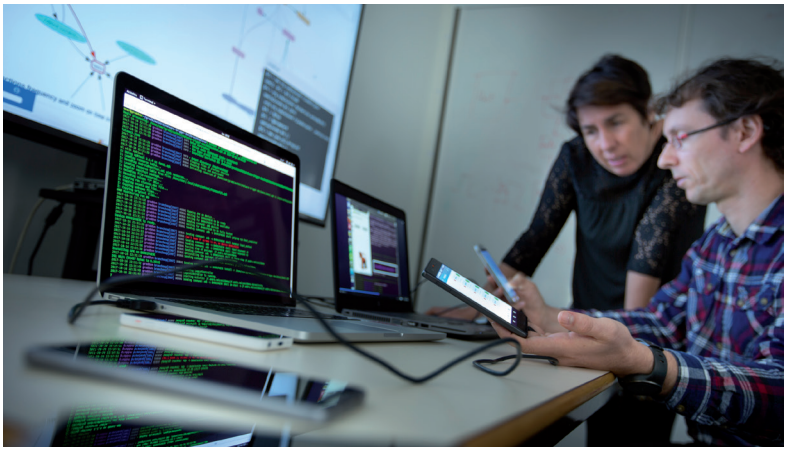
→ ***La seguridad y la privacidad están estrechamente relacionadas.*** El ransomware WannaCry no intentó exfiltrar los datos de los usuarios, pero podría haberlo hecho. El atacante tenía pleno acceso a los datos almacenados en las computadoras que eran objeto del ataque (por ejemplo, la base de datos de pacientes de un centro médico) y podría haber amenazado con revelar esta información sensible. Por tanto, es esencial que la seguridad y la privacidad se consideren conjuntamente en la fase de diseño para que, por ejemplo, las intrusiones maliciosas no pongan en peligro los datos. La seguridad por diseño, y más recientemente la privacidad por diseño, se han convertido en principios clave en el diseño de la seguridad.

→ ***La diversidad de motivaciones de los atacantes y la dificultad de atribución.*** Aunque WannaCry ha sido clasificado como ransomware, motivado por el deseo de ganar dinero, el ***malware*** NotPetya, que le sucedió rápidamente en junio de 2017, podría ser un "***malware*** patrocinado por el Estado que intentó disfrazarse de ***ransomware*** para enturbiar su atribución y potencialmente retrasar

las investigaciones"[8].

Estos ejemplos ponen de manifiesto la diversidad de las motivaciones de los atacantes y la dificultad –a veces, la imposibilidad– de atribuir un ataque.

➔ ***Detección y mitigación de ataques.*** Los ejemplos anteriores muestran que la seguridad es difícil de conseguir. Dado que el riesgo cero no existe, la detección temprana y la mitigación de los ataques son tan importantes como el intento de reducir el riesgo de ataques exitosos. En términos más generales, es probable que siempre haya vulnerabilidades en nuestros sistemas, a pesar de que los mecanismos de seguridad preventiva sean cada vez más eficientes. Las vulnerabilidades aparecen en todos los niveles de nuestros sistemas de información: aplicaciones, SO, firmware e incluso el hardware, como han ilustrado recientemente los ataques Meltdown y Spectre. A veces, las vulnerabilidades están presentes durante un tiempo (muy) largo en nuestros sistemas, y sólo podemos esperar que no se exploten antes de que se descubran. Diariamente se descubren nuevas vulnerabilidades[9] y en cualquier momento pueden aparecer nuevas formas de ataque. Es obligatorio detectar los ataques conocidos, pero también las nuevas formas de ataque, si queremos aumentar el nivel de seguridad de nuestros sistemas.



Análisis y detección de malware – © Inria / Foto C. Morel

➔ ***La seguridad tiene un costo.*** Es fácil entender que la seguridad puede ser cara, con costos adicionales para estudiar, implementar, configurar, gestionar y evolucionar las herramientas de seguridad. Pero la seguridad también puede

8. Obligations correspond to preconditions to fullfil to be permitted to read or write information. For example, a user can be authorized to sign a file if and only if this file has already been previously signed by another given user. Obligations are generally enforced at the application level.
9. https://www.cvedetails.com/browse-by-date.php

tener un costo operativo, lo que conduce a sistemas menos eficientes. Por ejemplo, para mitigar los ataques Spectre o Meltdown, puede ser necesario eliminar algunas técnicas de caché o desactivar la ejecución especulativa. Esta mitigación supondría una ralentización significativa y posiblemente inaceptable de la velocidad del procesador. Por lo tanto, en algunos casos, uno puede tener que aceptar un compromiso difícil entre la seguridad y la eficiencia.

Cada uno de estos temas es objeto de investigación activa y se presenta en este documento.

[Nota] La amenaza de la ciberseguridad

La amenaza de la ciberseguridad es real y grave. Sólo vemos la punta del iceberg: en la gran mayoría de los casos, incluso la existencia del ataque es una información crítica para las empresas o los estados que rara vez se da a conocer.

Para los expertos, la cuestión no es si los ciberataques a gran escala acabarán teniendo éxito, ya que la respuesta es definitivamente positiva, sino más bien: ¿estamos lo suficientemente preparados? Esto significa que, por supuesto, debemos reducir el riesgo de tales ataques mediante una mejor protección preventiva y reactiva, pero también aumentar nuestra ciber-resiliencia, incluyendo procedimientos preestablecidos para reducir los impactos catastróficos de los ataques exitosos, y una recuperación más rápida a un modo de operación seguro después de tales ataques.

1.4 Propiedades, servicios y mecanismos de seguridad

Propiedades de seguridad

La ciberseguridad consiste en garantizar tres propiedades básicas y esenciales de la información, los servicios y las infraestructuras informáticas, conocidas como la tríada CIA: Confidencialidad, Integridad y Disponibilidad *(Confidentiality, Integrity, and Availability)*. Así, garantizar la seguridad de un sistema de información significa impedir que una entidad no autorizada (usuario, proceso, servicio, máquina) acceda, altere o haga inaccesibles datos informáticos, los servicios informáticos o la infraestructura informática. Obsérvese que también podrían enumerarse otras propiedades, como la autenticidad (prueba del origen de la información), la privacidad o la protección contra las copias ilegales. Sin embargo, estas propiedades adicionales también pueden verse como casos particulares de estas tres propiedades básicas.

[Nota] Sobre las propiedades de la ciberseguridad

Confidencialidad: garantía de que la información se divulga sólo a personas, entidades o procesos autorizados.

Integridad: garantía de que el sistema (archivos de configuración, archivos ejecutables, etc.) o la información se modifican únicamente mediante una acción voluntaria y legítima, es decir, que el sistema o la información no han sido modificados accidental o deliberadamente.

Disponibilidad: garantía de que un sistema o información son accesibles en el momento oportuno para quienes necesitan utilizarlos.

Autenticidad: garantía de que un mensaje procede de la fuente que dice ser.

Privacidad: capacidad de los individuos de controlar sus datos personales y decidir qué revelar a quién y en qué condiciones. Por lo tanto, la privacidad puede definirse generalmente como el derecho de los individuos, grupos o instituciones a determinar por sí mismos cuándo, cómo y hasta qué punto se comunica la información sobre ellos a los demás.

Anonimato: confidencialidad de la identidad del usuario o de la entidad. Observamos que evitar la reidentificación mediante información lateral no es fácil, y que la indistinguibilidad, que garantiza que un atacante no pueda ver la diferencia entre un grupo de entidades, es también una propiedad importante vinculada a la privacidad. Nótese también que el anonimato tiene como objetivo ocultar quién realiza alguna acción, mientras que la privacidad total puede requerir también ocultar qué acciones se están realizando.

Política de seguridad: conjunto de reglas que especifican cómo se protegen los recursos sensibles y críticos, es decir, cómo se garantizan algunas o todas las propiedades anteriores.

Resiliencia: inicialmente definida como la capacidad de un sistema de volver a su estado original después de un ataque, la resiliencia se considera hoy en día como la capacidad de un sistema de prestar sus servicios de forma continua, incluso bajo ataque (es decir, la capacidad de tolerar ataques).

Servicios de seguridad

Alcanzar los objetivos de la ciberseguridad requiere aplicar contramedidas *físicas, organizativas y lógicas*. Aunque las medidas físicas (como la vigilancia o el control de los accesos a los edificios) y las medidas organizativas (como la definición precisa de la misión de un proveedor de servicios TI externo) son fundamentales, en este documento nos centramos en la seguridad lógica, es decir, en los servicios y mecanismos de hardware y software para garantizar las propiedades de confidencialidad, integridad y disponibilidad.

Un sistema informático seguro debe ofrecer *servicios preventivos* para impedir cualquier violación de estas propiedades, servicios de detección para identificar

cualquier intento exitoso de violar estas propiedades y servicios de reacción para desplegar contramedidas nuevas o mejoradas en caso de que una filtración o violación sea exitosa. Efectivamente, aunque el objetivo de la ciberseguridad es proteger un sistema informático contra los ataques, también hay que asumir que algunos de los ataques tendrán éxito. Por lo tanto, la ciberseguridad también se ocupa de la detección de intrusiones y de las respuestas a los ataques.

La prevención implica, en primer lugar, definir con precisión qué entidad puede acceder a qué información y de qué manera: hay que definir los permisos, las prohibiciones o las obligaciones[10] de leer o escribir información.

Esto constituye la llamada política de seguridad. La prevención puede tener lugar incluso antes de la definición de una política. De hecho, una buena ingeniería de software consiste en detectar con antelación las vulnerabilidades del código fuente y binario que podrían ser explotadas para violar las propiedades de seguridad: este es el principio de seguridad por diseño. Incluso antes, también podemos demostrar que una propiedad determinada está garantizada por el software: esto es la seguridad formalmente demostrada.

La política de seguridad se aplica concretamente a través de los servicios de seguridad. Dependiendo de la política y del contexto, se pueden ofrecer los siguientes servicios: identificación y autenticación de entidades, control del acceso a la información por parte de estas entidades, control de los flujos de información dentro del sistema, detección de los intentos de explotar las posibles vulnerabilidades del sistema (detección de intrusiones, detección de virus) y respuestas a estos intentos (reacción).

Existe un objetivo aún más ambicioso al que se puede aspirar: la capacidad de un sistema informático de obtener el resultado previsto a pesar de los acontecimientos cibernéticos adversos. En otras palabras, los sistemas informáticos tolerarían los ataques, una capacidad que suele denominarse ciber-resiliencia. Desde un nivel superior, una entidad (Estado, empresa, organización, etc.) podría estar más preocupada por la ciber-resiliencia, que se trata del objetivo principal que se quiere alcanzar, que por la ciberseguridad, que es un conjunto de técnicas desplegadas que el usuario final no tiene por qué ver.

La ciber-resiliencia[11], que es la capacidad de tolerar ataques, tiene por supuesto muchas similitudes con la tolerancia a fallos, que se ocupa de los fallos peligrosos del hardware o del software. Aunque las hipótesis de seguridad y de prevención son bastante diferentes, ya que los atacantes no siguen las reglas sino que buscan continuamente nuevas filtraciones y violaciones, los mecanismos propuestos para tolerar los fallos pueden adaptarse para tolerar los ataques.

10. Las obligaciones corresponden a las condiciones previas que hay que cumplir para poder leer o escribir información. Por ejemplo, un usuario puede obtener autorización para firmar un archivo si y sólo si este archivo ya ha sido firmado previamente.
11. Fuente https://en.wikipedia.org/wiki/Cyber_Resilience

Así, algunos principios básicos de la ciber-resiliencia incluyen la replicación de datos y las copias de seguridad, que están bien consolidadas desde hace tiempo en la comunidad de bases de datos. Además, la replicación debe utilizarse en el contexto de un sistema distribuido, para evitar tener un único punto de fallo. Aunque la ciber-resiliencia es de gran importancia, muchas técnicas para lograrla recuerdan a otros campos (por ejemplo, la seguridad), y no se detallarán en el resto de este documento; otras son completamente relevantes para el campo de la seguridad (por ejemplo, la mitigación de DDoS) y se discutirán en las secciones correspondientes.

Mecanismos de seguridad

Los servicios de seguridad se basan en mecanismos implementados en varios niveles de los sistemas e infraestructuras de información, como el hardware, el *firmware*, los sistemas operativos, las capas de red, los hipervisores y las aplicaciones. La criptografía es, por supuesto, un elemento fundamental en muchos casos: el estudio de las primitivas criptográficas y su uso en los intercambios entre máquinas (a través de protocolos criptográficos) son, por tanto, dos aspectos esenciales de la seguridad digital.

1.5 Aspectos jurídicos

1.5.1 Reglamento europeo de seguridad

La estrategia de ciberseguridad de la Unión Europea[12] se basa en varios instrumentos, con el objetivo de mejorar la ciber-resiliencia y la respuesta europea, preservando al mismo tiempo para cada nación un nivel de capacidad soberana para controlar los principales componentes de su estrategia de defensa nacional.

La Agencia de la Unión Europea para la Ciberseguridad (ENISA)[13] , creada en 2004, es un actor importante del panorama europeo de la ciberseguridad. Actualmente se está debatiendo una ampliación significativa de sus misiones, con el fin de convertirla en la interfaz privilegiada entre los Estados miembros, incluyendo el apoyo a la aplicación y el funcionamiento de las directivas de ciberseguridad.

A mediados de 2016, la Unión Europea adoptó la Directiva sobre la seguridad de las redes y los sistemas de información (conocida como Directiva NIS)[14] , para su aplicación a mediados de 2018. Esta directiva se centra tanto en los proveedores de servicios digitales (PSD) como en los operadores de servicios esenciales (OES por sus siglas en inglés). Tanto los PSD como los OES son responsables de

12. https://ec.europa.eu/digital-single-market/en/cyber-security
13. Inicialmente llamada Agencia Europea de Seguridad de las Redes y de la Información, véase https://www.enisa.europa.eu/
14. https://ec.europa.eu/digital-single-market/en/network-and-information-security-nis-directive

notificar los incidentes de seguridad, incluso si los servicios son gestionados por empresas no europeas o si la gestión del sistema de información se subcontrata a terceros. También se exige a los PSD y a los OES que proporcionen información que permita una evaluación en profundidad de la seguridad de su sistema de información y sus políticas. Por último, los Estados miembros deben identificar los organismos encargados de recoger y procesar los incidentes de seguridad, además de una autoridad nacional competente (por ejemplo, la ANSSI en Francia).

Asimismo, es necesario promover el desarrollo de productos y servicios "seguros por diseño" en toda Europa. Para lograr este objetivo, la Unión Europea propone establecer un marco europeo de certificación de la seguridad, capaz de emitir certificaciones de seguridad y etiquetas de productos y servicios a nivel europeo. Se trata de una tarea compleja, ya que, aunque existe una gran variedad de sistemas de certificación de seguridad en el mundo, no existe una solución unificada o combinada. Llegar a entender lo que se necesita para que todo sea sistemáticamente seguro es una tarea compleja. La Organización Europea de Ciberseguridad (ECSO)[15] colabora con la Comisión Europea para definir una propuesta de marco europeo de certificación de seguridad. La ECSO ha publicado un estado del arte de las normas de ciberseguridad industrial existentes para diversos ámbitos de actividad, y está trabajando en un enfoque denominado meta-esquema que abarca muchos esquemas de certificación existentes, evaluando el nivel de confidencia proporcionado por los esquemas individuales y asignándolos a un conjunto armonizado de niveles de seguridad[16].

Dado que el riesgo cero no existe, la Comisión Europea también ha publicado un proyecto oficial, *denominado Recomendación de la Comisión de 13.9.2017 sobre la Respuesta Coordinada a Incidentes y Crisis de Ciberseguridad a Gran Escala*[17]. Esta recomendación establece los objetivos y modos de cooperación entre los Estados miembros y las instituciones europeas a la hora de responder a dichos incidentes y crisis.

La Unión Europea está a la vanguardia de la privacidad y la protección de datos, con el nuevo reglamento RGPD y el reglamento ePrivacy que lo complementará –véase el apartado 5.1.2, donde se tratarán estos aspectos.

1.5.2 Análisis forense

En términos generales, el análisis forense está relacionado con los métodos científicos para identificar a los autores de un delito mediante el examen de objetos o sustancias implicadas en el mismo. En el contexto de la ciberseguridad, el análisis forense está relacionado con la explicación de un ciberdelito, basado en

15. https://www.ecs-org.eu
16. Los documentos elaborados por el ECSO WG1 están disponibles en http://www.ecs-org.eu/working-groups/wg1-standardisation-certification-labelling-and-supply-chain-management
17. http://ec.europa.eu/transparency/regdoc/rep/3/2017/EN/C-2017-6100-F1-EN-MAIN-PART-1.PDF

el análisis de la información o los rastros dejados por el atacante en los sistemas informáticos utilizados o atacados.

1.5.3 Vigilancia y seguridad

Con el aumento de la amenaza terrorista, hemos sido testigos, en varios países, del despliegue de sistemas de vigilancia masiva destinados a ayudar a combatir el terrorismo. En Francia, esto tomó la forma de la "*Loi relative au renseignement*"[18].En concreto, esta ley exige el despliegue de cajas negras en los proveedores de servicios de Internet (ISPs) franceses tanto para recopilar información de conexión de objetivos específicos previamente identificados en tiempo real, como para analizar la información de conexión de los abonados a los PSI con el fin de identificar a los posibles sospechosos mediante un proceso automático (cuyos detalles no se conocen públicamente). La re-identificación del abonado requiere una decisión oficial del Primer Ministro (o un delegado).

Estas leyes ponen de manifiesto la tensión entre la seguridad pública y la privacidad. También han sido criticadas por su costo económico y su potencial ineficacia, en particular cuando se enfrentan a la "paradoja del falso positivo."[19] El riesgo es que "más falsos positivos no harán más que sobrecargar las tecnologías, provocando así aún más trabajo para los agentes de inteligencia de señales, que ya están sobrecargados."

Se ha acuñado el término "*mass dataveillance*" para las prácticas en las que los gobiernos o las organizaciones gubernamentales llevan a cabo una vigilancia y una recopilación de datos a escala nacional (o mayor). Esto se opone a la "*personal dataveillance*", que se centra en un individuo de (supuesto) interés.

Como reacción a esta evolución de la vigilancia (y en particular a las revelaciones de E. Snowden), el IETF ha considerado que "la vigilancia omnipresente es un ataque" en el RFC 7258[20] y que los protocolos del IETF deberían mitigarla. El cifrado por defecto es una de las iniciativas del IETF (apartado 2.2). La vigilancia y la ciber-defensa son temas complejos por naturaleza.

La Alpine Multidisciplinary NEtwork on CYber Security studies (AMNECYS)[21] es un ejemplo de iniciativa multidisciplinaria que ha reunido a varios equipos de investigación para contribuir a esta compleja cuestión.

1.6 Cuestiones de soberanía

Dado que la mayoría de las infraestructuras críticas están ahora controladas

18. Loi no 2015-912 du 24 juillet 2015 relative au renseignement, Legifrance. https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000030931899
19. https://hal.archives-ouvertes.fr/hal-01157921/document
20. https://www.rfc-editor.org/rfc/pdfrfc/rfc7258.txt.pdf
21. http://amnecys.inria.fr/

por computadoras, a menudo conectadas a Internet, la protección de las infraestructuras requiere la protección de los sistemas y redes informáticos. De ahí que la ciberseguridad sea una cuestión de soberanía para los Estados y la UE. Por lo tanto, los Estados y la UE deben ser capaces de comprender los riesgos y las amenazas. Esto requiere los más altos conocimientos científicos, y sólo puede mantenerse a largo plazo si se lleva a cabo una investigación avanzada en todos los ámbitos de la ciberseguridad. No solo debemos contar con los mejores expertos, sino que debemos tenerlos en número suficiente para cubrir las crecientes necesidades (véase el apartado 2.3.3). Además, también necesitamos expertos de niveles intermedios e inferiores para poder aplicar correctamente las políticas de seguridad.

Los Estados también deben tener la capacidad de actuar. Esto requiere un control suficiente sobre la infraestructura digital y toda la cadena de ciberseguridad, ya que la seguridad del conjunto depende de la seguridad del eslabón más débil. Esto significa controlar el software y el hardware utilizados en las infraestructuras críticas, para que puedan ser analizados y certificados como libres de errores y puertas traseras. Esto también implica el control del almacenamiento de datos.

El hardware es uno de los eslabones más débiles, ya que Francia y Europa ya no tienen la capacidad de diseñar y producir su propio hardware. En consecuencia, ya se ha renunciado a cierta forma de soberanía. De hecho, es muy posible que los dispositivos de hardware estén equipados con puertas traseras o funciones ocultas que permitan, por ejemplo, a un organismo gubernamental o a una empresa espiar el tráfico de Internet o impedir el funcionamiento de un determinado servicio.

Es más, la naturaleza digital y, por tanto, desmaterializada de la ciberseguridad, hace que la soberanía en materia de ciberseguridad sea diferente de otras formas de soberanía, como la defensa. Mientras que esta última es un privilegio de los Estados u organizaciones supranacionales, la primera puede aplicarse a escalas más pequeñas. Muchas entidades (ciudadanos, empresas, asociaciones, etc.) pueden reclamar cierto grado de soberanía sobre la seguridad de sus propios datos, sistemas informáticos y redes. Una consecuencia de la digitalización es el potencial traspaso de algunas de las soberanías estatales tradicionales a otras entidades: el registro de la propiedad impulsado por el blockchain, la acuñación de dinero con monedas digitales, o los servicios de identificación ciudadana[22], etc. Estos distintos niveles de soberanía no se excluyen, sino que se complementan, dejando la soberanía de cada tipo de datos en el nivel más adecuado. Esta capacidad de descentralización de la ciberseguridad no debe poner en peligro las soberanías de los Estados. Al contrario, es una oportunidad que debe ser aprovechada, dejando cierta autonomía a las diferentes entidades dentro de ciertos límites establecidos por incentivos, reglamentos y leyes.

22 Por ejemplo, SecureIdentity https://secureidentity.co.uk/

2

Conocer, comprender y modelar las amenazas

Hay muchos tipos de ataques contra los sistemas de información. Por lo tanto, las amenazas son numerosas. Los ataques pueden dirigirse al hardware, a la red, al sistema o a las aplicaciones (muy a menudo a través de las acciones maliciosas de un malware), o incluso a los propios usuarios (ingeniería social, *phishing*). El atacante puede ser un *insider* o un *outsider*.

En este capítulo se presentan los trabajos realizados con el objetivo de profundizar en el conocimiento de estas amenazas y ataques. Estos trabajos se basan en varios modelos de ataque que definen el conocimiento del atacante sobre el sistema a atacar y las acciones que puede realizar sobre él. La sección dedicada al factor humano también se centra en dos aspectos cruciales de la seguridad: la usabilidad y la educación.

Tenga en cuenta que el criptoanálisis tiene como objetivo comprender las amenazas contra las primitivas criptográficas existentes, con el fin de adelantarse a los adversarios maliciosos. El criptoanálisis es, por tanto, la base de la confianza que se puede depositar en estas primitivas, y el conocimiento del criptoanálisis más avanzado puede considerarse la columna vertebral para el diseño de primitivas seguras. Por tanto, el criptoanálisis en sí mismo no es una amenaza, sino una forma de lograr un mejor nivel de seguridad. Se presenta un análisis del mismo en el apartado 3.1.2. El análisis y la detección de *malware* se presentan en el apartado 4.5.1.

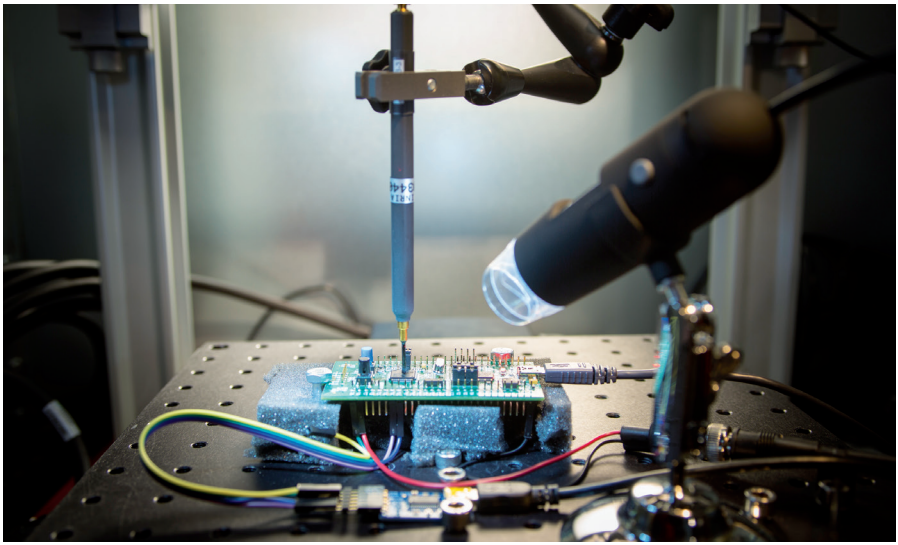
2.1 Ataques al hardware

[Resumen]

Los ataques físicos contra el hardware son una amenaza real, incluso para las implementaciones de algoritmos cuya seguridad se ha demostrado matemáticamente. Los ataques por observación y los ataques por perturbación son dos clases comunes de ataques que requieren un acceso físico al dispositivo. Más recientemente, el hardware también ha sido atacado a través del software. Esta forma de ataque es más peligrosa porque no requiere necesariamente el acceso físico al dispositivo atacado. Hoy en día, un escenario de ataque grave pero lamentablemente posible es un ataque de hardware desencadenado por una aplicación JavaScript incrustada en una página web.

Cuando se considera la seguridad de un algoritmo, los desarrolladores disponen de numerosas herramientas matemáticas para evaluarla. Lamentablemente, esas herramientas no pueden tener en cuenta la interacción de la unidad de computación con su entorno físico. Los ataques físicos son una amenaza real, incluso para los algoritmos cuya seguridad se ha demostrado matemáticamente. Estos ataques se pueden clasificar como ataques por observación, ataques por perturbación y un nuevo campo conocido como ataques de software dirigido al

hardware. Los dos primeros suponen el modelo de atacante interno, es decir, el dispositivo está bajo el control del atacante, mientras que el último supone el modelo externo. El modelo externo requiere menos hipótesis para el atacante y, por tanto, puede considerarse más peligroso.



Estudiando ataques de fallo – © Inria / Foto C. Morel

Ataques de observación

Los análisis de canal lateral (SCA) son ataques físicos basados en la observación del comportamiento del circuito durante un cálculo. Aprovechan el hecho de que algunas magnitudes físicas dependen de valores intermedios del cálculo en el dispositivo. Se trata de las llamadas fugas de información. Las fugas más clásicas son la sincronización de información, el consumo de energía, y las emisiones electromagnéticas (EM).

Los SCA son amenazas para todos los cripto-sistemas estándar, como el Estándar de Cifrado de Datos (DES), el Estándar de Cifrado Avanzado (AES), el criptosistema RSA, la Criptografía de Curva Elíptica (ECC) y para aplicaciones críticas que no utilizan criptografía, por ejemplo, la verificación de PIN. Los SCA también pueden ser utilizados para realizar ingeniería inversa de algoritmos.

Ataques de perturbación

Los ataques de fallo son una clase bien conocida de ataques físicos en los que un dispositivo sufre una modificación de parámetros físicos para obtener un comportamiento incorrecto. Los métodos de inyección de fallos más clásicos son

los fallos de potencia, los fallos de reloj, los pulsos láser y los pulsos electromagnéticos. Los ataques de fallo han demostrado ser extremadamente eficientes contra la criptografía, por ejemplo, el ataque Bellcore permite que cualquier fallo, en el momento correcto, en una firma RSA-CRT recupere el secreto. Qué fallo se puede conseguir y cuál es el modelo de fallo es un área activa de investigación.

Esta clase de ataque depende del chip, es decir, lo que se ha aprendido de un SoC no es válido para otros chips aunque sean muy similares (mismo núcleo). El éxito del ataque depende esencialmente de la configuración del experimento debido a la gran cantidad de parámetros (tipo de sonda electromagnética, distancia de la sonda en el circuito, forma y amplitud del pulso electromagnético, etc.). Otro reto en esta clase de ataques es la observabilidad del efecto. Para entender el efecto preciso, hay que explorar el estado interno del chip, que a menudo no está disponible. La mayoría de las contramedidas están relacionadas con la redundancia temporal o espacial. El costo de dicha redundancia no es asequible para los dispositivos de gama baja. La investigación se centra en la redundancia ligera para garantizar la integridad de la ejecución.

Ataques de software dirigidos al hardware

Además de los ataques de software contra el software y los ataques físicos dirigidos contra el hardware, en la segunda mitad de la década de 2000 aparecieron ataques de software contra componentes de hardware. Por ejemplo, el ataque Rowhammer tiene como objetivo cambiar bits de memoria (*flipping memory bits*) mientras lee y escribe otra celda. El modelo de atacante interno pasa a uno externo cuando se utiliza un programa JavaScript ejecutado en un navegador para realizar este ataque de forma remota. Recientemente se ha demostrado que es efectivo cuando se aplica en discos SSD (tecnología NAND flash).

La perturbación también puede generarse en SoC multinúcleo utilizando el *Dynamic Voltage and Frequency Scaling* (DVFS), es decir, la técnica de gestión energética que ahorra energía regulando la frecuencia y el voltaje de los núcleos del procesador. Se ha demostrado que una mala configuración de estos dos parámetros puede utilizarse para inducir fallos en el hardware. Al estar cada núcleo controlado individualmente, un núcleo puede inyectar un fallo en otro núcleo. Aunque todavía no se ha demostrado, este ataque podría realizarse desde un navegador.

Los ataques basados en software contra el hardware permiten burlar los mecanismos de seguridad implementados a nivel de software. Es más, las protecciones del software consideran que el hardware funciona correctamente, "simplemente" ejecutando instrucciones para producir un resultado. Por supuesto, esto no es tan fácil y los errores que pueden ser explotados por los atacantes también pueden

ocurrir a nivel de hardware.

En términos más generales, el enfoque tradicional de la informática y las tecnologías, que añade constantemente nuevos niveles de abstracción cada vez más potentes, lleva naturalmente, al proponer un mecanismo de seguridad a un nivel de abstracción determinado, a considerar que las capas inferiores son correctas y seguras. Sin embargo, esto no es así; por eso los atacantes han tenido una tendencia estos últimos años a tener como objetivo capas cada vez menos abstractas, atacando sucesivamente a través de software las aplicaciones, el SO, su kernel, el firmware, y ahora el hardware.

Estos ataques de capa inferior (*low-layer*) explotan normalmente los fallos de los mecanismos de optimización implementados en sistemas operativos y procesadores modernos, como las memorias caché, la predicción de ramas o la ejecución especulativa. De hecho, estas optimizaciones crean diferencias en el tiempo de ejecución del programa, revelando así información secreta. Por ejemplo, el reciente ataque Spectre[23] explota la predicción de ramas y la ejecución especulativa y exfiltra información a través de un canal encubierto basado en el acceso a la caché. Para mitigar este ataque, se podría refrescar las celdas (leer y reescribir sus valores) periódicamente. Por supuesto, esta solución tendría como consecuencia limitaciones de rendimiento, ya que otras operaciones de lectura solicitadas por los programas no serían posibles durante los refrescos. De forma más general, la protección contra este tipo de ataques implicaría la limitación, si no la eliminación completa, de ciertas optimizaciones, por supuesto a costa de un menor rendimiento.

El ataque Rowhammer mencionado anteriormente, es un ataque software que en realidad explota una propiedad física de la materia. Cada celda de la DRAM está compuesta por un condensador y un transistor que implementan eléctricamente un bit de información. Al acceder repetidamente a las celdas, la carga de éstas se escapa e interactúa eléctricamente con la carga de otras celdas vecinas. Es entonces posible cambiar el valor de una celda (y por lo tanto violar la integridad de esta celda) sin haber accedido nunca a ella. En este caso, la protección contra el ataque debería ser física: por ejemplo, se podría pensar en limitar la reducción de la superficie del componente, aunque el coste sería, por supuesto, muy importante.

Obsérvese que estos ataques no son fáciles de detectar, ya que no dejan ningún rastro a nivel del sistema operativo o de la aplicación.

Por último, es complicado saber si los ataques de este tipo ya se han utilizado de forma real. En el momento de redactar este documento, parece mucho más sencillo utilizar ataques más clásicos contra el software o contra los usuarios (ingeniería social).

23. https://spectreattack.com

[Desafío de investigación] Ataques de software dirigidos al hardware

Los ataques contra los sistemas de información no suelen afectar a la capa de hardware, sino que explotan una vulnerabilidad del software. Sin embargo, ataques recientes, como Rowhammer, Spectre o Meltdown, han demostrado que los ataques implementados en software pueden explotar las optimizaciones de rendimiento del hardware. Este nuevo tipo de ataque es especialmente peligroso, ya que hace posible los ataques al hardware a distancia, a diferencia de los ataques clásicos de canal lateral. Todavía no está del todo claro cómo se pueden "industrializar" los ataques actuales de prueba de concepto, pero abren el camino a una nueva clase de ataques graves. Por lo tanto, es necesario comprender mejor cómo podrían desplegarse estos ataques, proponer una tipología clara de este nuevo tipo de ataque y proponer contramedidas, tanto a nivel de hardware como de software. Esta tarea requiere experiencia en los niveles de hardware, firmware y sistema operativo. Las contramedidas también pueden ser difíciles de diseñar, ya que pueden requerir la revisión de optimizaciones cruciales utilizadas durante años, como la ejecución especulativa.

[Equipos Inria] Ataques al hardware

↗ El equipo **CAIRN** trabaja en la arquitectura de computación eficiente. No abordan explícitamente los problemas de seguridad relacionados con este proceso de optimización, pero saben cómo puede ayudar la arquitectura del SoC.

↗ El equipo **CIDRE** analiza el impacto de los fallos EM en la pila (stack) de software de los SoC modernos, pero también en los dispositivos de gama baja. El equipo también utiliza la observación de circuitos para optimizar el proceso de inyección de fallos. El objetivo es evaluar la posibilidad de pasar de una plataforma de gama alta a una de bajo costo con el mismo resultado. El equipo caracteriza el impacto del fallo durante el proceso de arranque seguro para identificar posibles vulnerabilidades. En colaboración con el equipo **PACAP**, verificaron la solidez de una contramedida que utiliza un compilador específico para generar código tolerante a los fallos. Para resistir a los ataques de canal lateral, evaluaron soluciones de compilación sobre la marcha para aumentar el número de rastros que un atacante tiene que capturar para extraer el secreto. También han propuesto un modelo de ataque contra las implementaciones de verificación de PIN.

↗ El equipo **PACAP** evalúa la posibilidad de erradicar las fugas en un código en tiempo de compilación gracias a la anotación de código (compilador que genera implementaciones resistentes a los ataques de análisis de canal lateral). Su solución se evalúa en las instalaciones del LHS de Rennes.

↗ El equipo **TAMIS** desarrolla nuevos diferenciadores de canales laterales basados en técnicas de aprendizaje automático gracias a su conocimiento preciso de las mediciones de fugas subyacentes y al modelado de la información sensible.

2.2 Amenazas de la red

[Resumen]
A nivel de red, existen muchos ejemplos de ataques. He aquí dos ejemplos dirigidos a la Internet. Encontrar una ruta para cada paquete enviado en Internet, independientemente de su origen y su destino, es un servicio clave conocido como "enrutamiento"a: atacar este servicio de red básico y esencial puede, por ejemplo, aislar a todo un país o, al contrario, redirigir todo el tráfico de un país a través de un punto de vigilancia. Otro servicio de red crucial, el DNS, traduce los nombres de host legibles en direcciones IP. Un ataque contra este servicio puede redirigir a un usuario a un sitio web bancario falso para robar sus credenciales. Ya existe una extensión segura del DNS, llamada DNSSEC, pero su implementación llevará tiempo y no resolverá todos los problemas, en particular aquellos relacionados con la privacidad.

a. Véase por ejemplo esta introducción simple: https://interstices.info/internet-le-conglomerat-des-reseaux/

Cualquier tipo de red puede ser atacada, aprovechando sus características. Aquí nos centramos en Internet y en algunas de sus especificidades: nombre de dominio, enrutamiento y carga útil potencialmente no cifrada.

La Internet es un conjunto complejo de una cantidad extremadamente grande de dispositivos, desde máquinas o dispositivos de usuarios hasta routers, conectados por una enorme gama de tecnologías de red inalámbricas y por cable. Su funcionamiento requiere una amplia gama de recursos de información, protocolos y servicios, desde bases de datos de enrutamiento de bajo nivel, políticas de reenvío, protocolos de mapeo de direcciones MAC/IP/ host-name de bajo nivel (es decir, los protocolos ARP/NDP/SEND y DNS/DNSSEC), o tecnologías específicas de la capa de enlace (por ejemplo, para crear y gestionar una LAN virtual) hasta servicios de alto nivel, como los servicios web. Esta complejidad intrínseca de la Internet constituye muchas facetas que están sujetas a amenazas. En este artículo, consideramos un pequeño subgrupo de estas amenazas y analizamos las tendencias para mitigarlas, tanto desde el punto de vista académico como de la estandarización, por ejemplo, como hace el Grupo de Trabajo de Ingeniería de Internet (IETF)[24].

24. https://ietf.org

Ataques contra el sistema de nombres de dominio (DNS)

El DNS es un sistema de nomenclatura jerárquico y descentralizado para Internet, cuyos objetivos principales son la escalabilidad y la flexibilidad. El DNS se utiliza para la resolución de direcciones, es decir, el mapeo de nombres hosts a direcciones IP (por ejemplo, "www.example.com" se resuelve en la dirección IPv4 "1.2.3.4"), así como el mapeo inverso. También lo utilizan servicios como el correo electrónico (los registros DNS permiten buscar servidores de correo) y los hosts de correo electrónico de lista negra.

Un ataque típico contra el DNS consiste en inundar un servidor DNS con un gran número de consultas, lo que conduce a una denegación de servicio, es decir, el servidor no puede manejar la carga y por lo tanto no responde a las consultas legítimas. Un ataque más sutil consiste en envenenar o hacer spoofing de una caché DNS. En efecto, cuando un sistema consulta un servidor DNS y recibe una dirección IP como respuesta, guarda esta información en una caché local durante un periodo de tiempo determinado, de manera que el sistema puede responder a una nueva consulta similar sin tener que recuperar nuevamente la información del servidor. Si la caché se ve comprometida, cualquiera que la utilice puede ser conducido erróneamente a un sitio fraudulento.

Al ser una de las piedras angulares de Internet, la seguridad del DNS es esencial y en 2005 se añadieron servicios de seguridad bajo el nombre de "Extensiones de seguridad de nombre de dominio" (o DNSSEC). DNSSEC permite que cualquier host confíe en los resultados de la resolución de la dirección, la dirección IP (o cualquier información devuelta por una consulta DNS como los servidores de correo).

Sin embargo, el DNSSEC no se ocupa de los requisitos de confidencialidad, como se verá en el apartado 5.3.6.

Ataques contra el enrutamiento interdominio del Protocolo de Puerta de Enlace de Frontera (BGP)

Cada datagrama IP (o paquete) necesita ser "enrutado" a través de Internet, desde su origen hasta su destino (a veces múltiples destinos). Esta operación la realizan los routers salto a salto: para cada paquete entrante, un router encuentra una ruta y reenvía el paquete al siguiente router hasta que llega a su destino. El objetivo de los protocolos de enrutamiento es encontrar una ruta dentro de un router, y estos protocolos de enrutamiento aprovechan una base de datos distribuida que contiene información de enrutamiento y alcanzabilidad. Existen dos tipos de protocolos de enrutamiento: algunos protocolos están pensados para operar dentro de sistemas autónomos (AS), controlados por una única organización (por ejemplo, una universidad), mientras que otros están pensados para operar a nivel de interconexión, entre sistemas autónomos, es decir, a nivel de la red troncal de Internet. El Border Gateway Protocol (BGP) es el protocolo que se utiliza actualmente en Internet para el intercambio de información de

enrutamiento y alcanzabilidad entre sistemas autónomos. Por lo tanto, es de suma importancia, ya que cualquier comportamiento erróneo, tal vez causado por un ataque, puede aislar a todo un país de Internet, o redirigir todo el tráfico de un país determinado a través de un punto de vigilancia[25]. BGP lleva mucho tiempo sufriendo debilidades de seguridad. Por ejemplo, un atacante puede falsificar una respuesta BGP que le permita secuestrar más tráfico. Para resolver estas debilidades es necesario:

- establecer una infraestructura de clave pública dedicada a distribuir certificados que puedan ser verificados por cualquier enrutador BGP. Esta es la función de la Infraestructura de Clave Pública de Recursos (RPKI), gestionada por los distintos registros de Internet (IANA y Registros Regionales de Internet);
- aprovechar esta RPKI para emitir certificados, llamados Route Origination Authorization (ROA), que confirman que un AS controla ciertos rangos de direcciones IP y está autorizado a originar anuncios de rutas para estos rangos de direcciones IP.

Aunque estos mecanismos son necesarios, una ROA por sí misma no evita que un atacante (por ejemplo, un enrutador BGP malicioso) falsifique o propague anuncios de ruta maliciosos. El problema subyacente de validar una ruta de AS para un destino específico es complejo y multidimensional: requiere comprobar la validez del AS, la vecindad del AS, la conformidad de la ruta del AS listada en el mensaje con la propagación del mensaje en sí, y la conformidad de la ruta del AS con las políticas de enrutamiento reales de cada AS. La extensión de seguridad de BGP, BGPsec (véase más adelante), pretende ofrecer algunas garantías desde este punto de vista.

El grupo de trabajo del IETF sobre enrutamiento seguro entre dominios (SIDR) ha especificado las extensiones de seguridad RPKI y BGP. Sin embargo, el despliegue lleva tiempo, especialmente con BGPsec que requiere actualizaciones importantes, y contar con un enrutamiento interdominio totalmente seguro sigue siendo un sueño lejano. Siendo el despliegue parcial la regla, varios trabajos académicos se centran en las consecuencias de los ataques que se originan en partes no seguras de Internet y en las técnicas para mitigarlos.

Otro aspecto de BGP es la importancia geoestratégica de la información de enrutamiento. Su análisis puede ayudar a identificar o comprender los ataques (por ejemplo, cuando un subconjunto de Internet se vuelve repentinamente inalcanzable) o las prácticas de vigilancia (por ejemplo, cuando un subconjunto del tráfico de Internet se redirige a través de un determinado dominio).

25. Aislar una zona geográfica también podría estar motivado por el deseo de desconectar un determinado número de los servidores implicados en los servicios distribuidos y facilitar así un ataque (por ejemplo, un servicio basado en blockchain).

Encriptación por defecto y mitigación de los ataques de vigilancia masiva

"Espiar es irresistible. Si hay información visible en el paquete, no hay forma de evitar que un nodo intermedio la vea. Así que la última defensa del modo extremo a extremo es el cifrado extremo a extremo".[CWSR02] Siguiendo este sabio consejo, se han llevado a cabo actividades de investigación y estandarización para facilitar el uso del cifrado dentro de Internet y para frustrar las escuchas pasivas en particular.

Este es el caso del *cifrado TCP*, por ejemplo, tal y como lo ha desarrollado el grupo de trabajo del IETF TCP Increased Security (tcpinc)[27]. La idea que subyace al cifrado TCP es diseñar extensiones TCP que puedan proporcionar un cifrado no autentificado y una protección de la integridad de los flujos TCP. En este caso, el mecanismo de intercambio de claves no autenticado permite a ambos extremos cifrar y comprobar la integridad de cada paquete TCP, de forma muy sencilla, sin depender de ningún servicio externo (por ejemplo, la infraestructura PKI), ni de la solicitud del usuario (como ocurre con SSL cuando se conecta a un nuevo host), ni de ninguna modificación de las aplicaciones para las que esta extensión es totalmente transparente. Dicha extensión, una vez desplegada de forma suficiente, permitirá que todos los flujos TCP estén cifrados por defecto. Sin embargo, como no hay autenticación extremo a extremo, existe el riesgo de ataques *Man-in-the-Middle*: el atacante se hace pasar por el nodo remoto y propone su propio material de clave que le permite descifrar, espiar y volver a encriptar todo el tráfico. Por esta razón, se considera como una seguridad de tipo "peor es nada".

La encriptación por defecto tiende a convertirse en la norma, como es el caso del nuevo protocolo de transporte QUIC, desarrollado actualmente en el grupo de trabajo[28] del IETF QUIC. Propuesto inicialmente por Google como sustituto de alto rendimiento de HTTP sobre conexiones TLS/TCP, este protocolo ya representa una parte significativa del tráfico de Internet (más del 30% del tráfico de salida de Google, o el 7% del tráfico de Internet, a finales de 2016). Entre muchas de las innovaciones, este protocolo proporciona por defecto comunicaciones seguras: los paquetes QUIC siempre se autentifican y la carga útil suele estar totalmente encriptada, lo que impide la vigilancia masiva y otras formas de ataque.

[Equipos Inria] Amenazas de la red

↗ El equipo **DATASPHERE** trabaja en el análisis de los datos BGP para identificar o comprender los ataques o las prácticas de vigilancia.

[CWSR02] D. Clark, J. Wroclawski, K. Sollins, and Braden R. Tussle In cyberspace: Defining tomorrow's internet. En proceedings of SIGCOMM, 2002.

27. https://datatracker.ietf.org/wg/tcpinc/about/
28. https://datatracker.ietf.org/wg/quic/about/

2.3 El factor humano

Como en muchos otros campos, hay un dicho muy conocido sobre la seguridad que dice que la principal amenaza está entre la silla y el teclado. Puede que este dicho sea exagerado, y como mínimo merezca un estudio más profundo, pero hay que reconocer que los usuarios son en ocasiones una fuente de problemas de seguridad. En primer lugar, pueden ser el objetivo del ataque (véase el apartado 2.3.1). Además, pueden intentar evitar el uso de los mecanismos de protección disponibles debido a la excesiva complejidad (real o percibida) de su utilización (véase el apartado 2.3.2). Por último, su nivel de educación y formación es a menudo insuficiente (véase el apartado 2.3.3): por tanto, no son conscientes de los riesgos reales o, por el contrario, los sobreestiman. En cualquiera de los dos casos, no saben qué mecanismos hay que utilizar y cuándo.

2.3.1 Ataques contra el usuario: ingeniería y phishing

[Resumen]
La ingeniería social tiene como objetivo convencer al usuario para que lleve a cabo una acción, como revelar una contraseña, ganando su confianza. Estrechamente relacionado, el phishing tiene como objetivo obtener información como contraseñas, números de tarjetas de crédito, etc. Suele basarse en campañas masivas de correo electrónico (spam) o en mensajes a través de otros medios de comunicación (chats, redes sociales) solicitando a las personas que proporcionen información sensible, ya sea respondiendo a un correo electrónico o conectándose a un sitio web.

La ingeniería social tiene como objetivo convencer a una persona para que realice una acción prohibida o delicada ganando su confianza. El atacante puede suplantar la identidad de una persona o puede poner como pretexto un asunto urgente falso. La ingeniería social no se limita estrictamente a Internet. Sin embargo, la Internet permite ampliar los ataques de phishing, uno de los principales exponentes de los ataques de ingeniería social. El phishing suele tener como objetivo obtener información como contraseñas, números de tarjetas de crédito, etc. Suele basarse en campañas masivas de correo electrónico (spam) o en mensajes a través de otros medios de comunicación (chats, redes sociales) para que la gente proporcione información sensible respondiendo al correo electrónico o conectándose a un sitio web. Al principio, las campañas de phishing eran bastante sencillas e ingenuas porque se enviaba el mismo correo electrónico de forma masiva sin ningún tipo de personalización. El *spear phishing* es más avanzado y aprovecha una mayor inteligencia social para hacer que la gente confíe en la legitimidad de la solicitud

que ha recibido. Para ello, la solicitud puede personalizarse en función del país, la ubicación o la empresa de las víctimas. En realidad, este tipo de información puede encontrarse fácilmente en Internet.

El FBI calcula que el dinero extorsionado por el phishing fue de unos 500 millones de dólares al año entre 2013 y 2016 [29] en base a las denuncias recogidas en EEUU. Aunque hay técnicas sofisticadas que pueden apoyar el phishing, como el envenenamiento de DNS para apoderarse de un sitio web legítimo, los atacantes suelen recurrir normalmente a técnicas de ingeniería social, que son una forma mucho más fácil de lograr el mismo fin.

Aunque la ingeniería social masiva e ingenua puede detectarse fácilmente, es difícil detectar el *spear phishing* incluso si se lleva a cabo a gran escala, debido a la personalización automática de cada correo electrónico enviado. La defensa se produce en varios niveles. En primer lugar, los ataques de phishing se basan en un canal de comunicación para llegar a las víctimas. Para las grandes campañas, los atacantes aprovechan las redes de bots. Por lo tanto, la lucha contra las redes de bots limita indirectamente los ataques de phishing. En segundo lugar, el análisis de los mensajes de los usuarios, el contenido en sí mismo y la correlación entre ellos, es otra opción. En tercer lugar, la mayoría de las veces, la víctima es redirigida a un sitio web. Detectar o impedir la existencia o el acceso a estos sitios web es también una contramedida eficaz. Tanto el segundo como el tercer enfoque son relevantes para la ingeniería social, ya que, en estos casos, el atacante necesita cierta inteligencia social para hacer que las víctimas confíen en la legitimidad del correo electrónico o del sitio web.

El análisis del contenido de los mensajes de correo electrónico es problemático debido a la preocupación en torno a la privacidad. Además, con el uso del cifrado, sólo los usuarios finales o sus servidores de correo electrónico pueden tener acceso al contenido, lo que hace imposible la correlación de múltiples correos electrónicos enviados a usuarios distintos. Esto significa que sólo es posible el análisis individual de los correos electrónicos. Por ello, el uso de listas negras de nombres de dominio o URLs está muy extendido. Además, bloquear el acceso a un solo sitio web de este tipo significa proteger a miles de usuarios de una sola vez. En realidad, el principal reto consiste en bloquear con antelación un sitio web o una URL de phishing e incluso predecirlo cuando sea posible. De hecho, el atacante necesita configurar su infraestructura antes de iniciar su campaña de forma efectiva. Por supuesto, los sitios web sólo pueden ser analizados una vez que se conoce una URL potencial. Por lo tanto, se necesitan técnicas más activas y reactivas para descubrir y verificar automáticamente los posibles sitios web de phishing.

29. https://www.ic3.gov/media/2017/170504.aspx#fn3

[Equipos Inria] El factor humano

↗ El equipo **RESIST** trabaja en la construcción de listas negras automáticas de posibles sitios web de *pishing*. Este trabajo incluye un enfoque reactivo y otro proactivo.

2.3.2 Mejora de la usabilidad del mecanismo de seguridad

[Resumen]

Una de las principales fuentes de fallos de seguridad informática siguen siendo los errores humanos. Uno de los principales motivos es que las interfaces de usuario de muchas aplicaciones o sistemas de software no suelen estar diseñadas teniendo en cuenta la seguridad. Una buena interfaz de usuario que tenga en cuenta la seguridad debe tener en cuenta que el usuario rara vez es un experto en seguridad; debe asegurarse siempre que el usuario sea consciente de las consecuencias de sus acciones y estar diseñada para evitar errores involuntarios que comprometan la seguridad. El diseño de un mecanismo de seguridad con buena usabilidad exige una investigación interdisciplinar con expertos en ciencias cognitivas.

Los errores humanos son una de las principales fuentes de fallos de seguridad informática. Si bien es fundamental aumentar la conciencia de los usuarios sobre los riesgos de seguridad, suponer que los usuarios están bien formados no es realista ni resulta suficiente. Una de las principales razones de los errores humanos es que las interfaces de usuario de los sistemas de software no están diseñadas teniendo en cuenta la seguridad, si bien los principios de diseño adecuados para otras aplicaciones sí se aplican. En efecto, la seguridad no es el objetivo principal del usuario, a diferencia de lo que ocurre con la navegación por la web o la compra de bienes y servicios. En el mejor de los casos, los usuarios quieren que la seguridad funcione sin que ellos tengan que realizar ninguna acción específica; a menudo es más fácil ignorar o eludir los mecanismos de seguridad si esto facilita la obtención de su objetivo principal. Un ejemplo clásico es el de un sitio web con un certificado incorrecto; el simple hecho de ignorar la advertencia permite al usuario seguir navegando, con el riesgo consiguiente de visitar un sitio web falso, de phishing. Otro ejemplo es la desactivación de todo el *firewall,* para asegurarse de que un determinado puerto utilizado por una aplicación específica no está bloqueado. A menudo, los usuarios no comprenden, o malinterpretan, las consecuencias de sus acciones. Por lo tanto, los mecanismos de seguridad deben diseñarse teniendo en cuenta la usabilidad. La interfaz de usuario debe tener en cuenta que el usuario no es un experto en seguridad (y no entiende, por ejemplo, los mecanismos subyacentes de control de acceso), asegurarse que el usuario es plenamente consciente de las consecuencias de sus acciones y evitar que cometa errores que comprometan la seguridad. El diseño de un mecanismo

de seguridad utilizable exige una investigación interdisciplinar con expertos en ciencias cognitivas.

[Equipos Inria] Una usabilidad a mejorar

↗ El equipo **CIDRE** trabaja con ergónomos y psicólogos para evaluar la forma en que el usuario percibe la importancia de un mensaje, con el fin de que el usuario sea plenamente consciente de su acción.

[Reto de investigación 2] Seguridad y usabilidad

Muy a menudo, cuando los usuarios solicitan un servicio, están dispuestos a sacrificar la seguridad, y a saltarse un mecanismo de seguridad molesto, si ese mecanismo les impide utilizar el servicio. Para evitar este problema, la seguridad debe ser lo más transparente posible. Aunque la transparencia total no siempre es posible, los servicios de seguridad deben ser lo más sencillos posible de utilizar. Hay que trabajar para proponer interfaces y mecanismos de seguridad que sean adecuados para los usuarios no expertos, que garanticen que el usuario es plenamente consciente de la consecuencia de sus acciones y que eviten que los usuarios cometan errores que comprometan la seguridad. El diseño de estos mecanismos de seguridad utilizables exige una investigación interdisciplinar que incluya, por lo general, a expertos en ciencias cognitivas.

2.3.3 La falta de educación y concientización

[Resumen]

A menudo se considera a los usuarios como el eslabón más débil de la cadena de seguridad, ya que a menudo no son conscientes de los problemas de seguridad y, por lo tanto, son fácilmente engañados incluso por ataques muy simples. Por eso es fundamental la educación y la concienciación sobre las "buenas prácticas" y la "estrategia de ciber-higiene" de cada usuario de computadora (en el trabajo o en casa). La gente joven debe ser iniciada en los conceptos básicos de la informática y la ciberseguridad. Cada actor profesional debería conocer los riesgos relacionados con la inteligencia económica y los ciber-ataques, y estar informado de las posibles defensas. Además, los administradores de sistemas y redes deben recibir formación periódicamente para estar al día sobre las amenazas más recientes y las soluciones para mitigarlas.

El sector necesita expertos en ciberseguridad, a pesar de que se enfrenta en la actualidad a una escasez de expertos en ciberseguridad a todos los niveles. Si bien existe una cantidad significativa de instituciones que en la actualidad ofrecen planes de estudios de ciberseguridad, hay que seguir invirtiendo en la formación de más expertos.

Cuando los usuarios (ciudadanos o partes interesadas) son afectados por la ciberseguridad o la falta de ella, tienen poca consciencia de lo que está pasando. Con demasiada frecuencia las víctimas de ciber-ataques son considerados por muchos como el eslabón débil de la cadena, ya que no son conscientes de los problemas de seguridad y, por lo tanto, son fácilmente engañados incluso por ataques tecnológicamente muy simples como el phishing o la ingeniería social. Al estar poco informados de la importancia de protegerse, pueden, por ejemplo, utilizar contraseñas demasiado débiles o demasiado predecibles. Por último, una determinada funcionalidad puede ser más importante para ellos que la seguridad, por lo que tienden a hacer clic sistemáticamente en un enlace web o a utilizar un certificado no firmado. Por tanto, un reto importante es hacer que las herramientas de seguridad sean más fáciles de usar (véase el apartado 2.3.2).

En términos más generales, la mayoría de los ataques exitosos aprovechan problemas de seguridad bien conocidos y una gran mayoría de los ciber-ataques son el resultado de malos hábitos cibernéticos dentro de las organizaciones de las víctimas.

En este contexto, la educación y la concientización de cada usuario son esenciales. Los usuarios deben ser educados sobre las "buenas prácticas" para situaciones domésticas y profesionales y deben ser capaces de aplicar una "estrategia de ciber-higiene" con el fin de reducir los riesgos de convertirse en víctima o de propagar un ataque.

Como primer paso, todo joven debería ser iniciado en los conceptos y herramientas básicos de seguridad, al mismo tiempo que en los fundamentos de la informática. Algunos de los temas incluyen la importancia de actualizar el software o el sistema operativo para evitar contar con demasiadas vulnerabilidades, la importancia de un antivirus, la definición de una buena contraseña o una firma electrónica.

Un segundo paso es educar a cada parte interesada profesional, en cada programa educativo, sea cual sea el ámbito, introduciendo buenas prácticas adicionales relativas al contexto profesional. Hay que explicar a cada profesional los riesgos de la inteligencia económica, enseñarle la separación entre los datos y aplicaciones profesionales y los personales, la partición de redes y a realizar copias de seguridad y planes de continuidad del negocio. En Francia, un buen ejemplo de lo que podría hacerse es el proyecto CyberEdu[30] de la ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information), que tiene como objetivo desarrollar material pedagógico que facilite la integración de la seguridad digital en la enseñanza superior para los no especialistas.

Un tercer paso se dirige principalmente a los equipos operativos en las pequeñas y medianas empresas. Estos equipos deben conocer el estado del arte en materia de ciberseguridad y, por tanto, deben seguir una formación periódica sobre los riesgos y amenazas recientes, sobre la puesta a punto de los sistemas, sobre el refuerzo de

30. https://www.ssi.gouv.fr/entreprise/formations/cyberedu/

la seguridad, sobre el mantenimiento de un estado de seguridad global y sobre la legislación y la reglamentación. Varias iniciativas siguen este camino: en Francia, la ANSSI ha elaborado la guía "40 medidas esenciales para una red sana" [31] dedicada a los responsables de la seguridad de los sistemas de información. A nivel de la UE, la "Red de Recursos Humanos de Ciberseguridad" de la ECSO pretende aumentar los niveles de concienciación a través de diversas iniciativas de ciber-higiene, y la ENISA publicó su propio documento[32].

Por último, la formación de los profesionales de la ciberseguridad es, por supuesto, una tarea fundamental. En la actualidad, el sector ya se enfrenta a una escasez de trabajadores cualificados en ciberseguridad, como demuestra el número de ofertas de empleo publicadas en la APEC[33] para puestos de ciberseguridad, que en Francia se han cuadruplicado, pasando de 315 ofertas a 1.133 ofertas[34] entre 2014 y 2016. La empresa de consultoría empresarial Frost & Sullivan pronosticó que, para el año 2020, faltarán 1,5 millones de profesionales en todo el mundo, según el estudio del año 2017 Global Information Security Workforce Study[35].

Solucionar este problema representa un esfuerzo a largo plazo en materia de educación y formación. La dificultad inherente es que la ciberseguridad requiere una formación muy sólida en informática, junto con conocimientos adicionales relativos al entorno de las amenazas, los conceptos y las herramientas de seguridad, y una buena comprensión del derecho, los factores humanos y la psicología, las ciencias sociales, la economía y la gestión de riesgos.

Un gran número de instituciones ahora proponen planes de estudios técnicos. En Francia, la ANSSI ha lanzado SecNumEdu[36], cuyo objetivo es certificar algunos planes de estudios para garantizar a los estudiantes y a los empresarios que la formación en el ámbito de la seguridad digital cumple los criterios definidos por la ANSSI en colaboración con la industria y las instituciones de enseñanza superior. Una primera lista de programas ya está redactada[37].

. Un trabajo similar ha sido realizado por la OTAN, principalmente para las universidades de los EE. UU. y el Reino Unido[38] y la ENISA para las universidades de la UE[39]. También se pueden citar, a nivel europeo, las Actividades Digitales en Materia de Ciberseguridad del EIT, que incluyen escuelas profesionales[40] y de magísteres[41].

31. https://www.ssi.gouv.fr/uploads/IMG/pdf/guide_hygiene_v1-2-1_en.pdf
32. https://www.enisa.europa.eu/publications/cyber-hygiene
33. https://www.apec.fr/
34. http://www.bretagne.bzh/upload/docs/application/pdf/2017-06/etude_apeccybersecuritebretagne.pdf
35. https://iamcybersafe.org/wp-content/uploads/2017/06/Europe-GISWS-Report.pdf
36. https://www.ssi.gouv.fr/entreprise/formations/secnumedu/
37. https://www.ssi.gouv.fr/particulier/formations/secnumedu/formations-labellisees-secnumedu/
38. https://digitalguardian.com/blog/cybersecurity-higher-education-top-cybersecurity-colleges-and-degrees
39. https://resilience.enisa.europa.eu/nis-platform/shared-documents/wg3-documents/cybersecurity-education-snapshot-for-workforce-development-in-the-eu/view
40. https://www.eitdigital.eu/eit-digital-academy/professional-school/
41. https://www.eitdigital.eu/eit-digital-academy/master-school/

El carácter interdisciplinario de la ciberseguridad puede requerir métodos y estrategias de enseñanza novedosos. En este caso, el aprendizaje práctico (y no solo teórico) es fundamental. Los programas también podrían aprovechar los MOOC y la formación profesional técnica específica, como los ejercicios de ciberdefensa (CDX) a gran escala. Un ejemplo de estos últimos es el Locked Shields 2017[42] de la OTAN.

2.3.4 Manipulación de los usuarios y la opinión pública

[Resumen]
Los datos y las tecnologías pueden utilizarse para influir en las opiniones o decisiones en línea, mediante el profiling y las noticias falsas bien dirigidas. Esto requiere soluciones técnicas para disuadir el profiling no deseado y permitir una fácil verificación de la información, así como la validación de su fuente.

Un problema grave, que en realidad está poco relacionado con la seguridad, pero que a menudo está vinculado a ella, al menos en la mente del público en general, es que los datos y las tecnologías también se utilizan para motivar, influir o moldear las opiniones o decisiones de las personas en línea. Una mejor comprensión de los comportamientos de los usuarios, combinado con la capacidad de construir perfiles psicológicos precisos, genera oportunidades para desarrollar técnicas que influyen en los usuarios en línea, moldeando sus opiniones. Estas tecnologías tratan de influir en el razonamiento y las decisiones del usuario manipulando sus "sesgos cognitivos", por ejemplo, sus emociones, su memoria o sus creencias.

Una forma particular de manipular la opinión pública son las noticias falsas. A nivel mundial, la desinformación a través de Internet se considera hoy en día un problema importante que requiere soluciones técnicas que permitan una fácil verificación de la información, así como la validación de sus fuentes. En el sector de los medios de comunicación, varias iniciativas recientes han comenzado a identificar, rastrear y desmentir las noticias falsas y las afirmaciones erróneas que circulan por Internet, por ejemplo creando un repositorio compartido de noticias falsas o listas de fuentes de información fiables y no fiables (*véase, por ejemplo, la base de datos Decodex*[43] *de Le Monde en Francia*).

42. https://ccdcoe.org/locked-shields-2017.html
43. http://www.lemonde.fr/verification/

[Equipos Inria] Manipulación de los usuarios y de la opinión pública

↗ El equipo **PRIVATICS** investiga cómo se utilizan los datos personales para manipular a la gente en Internet.

Inria, junto con socios de la industria de los medios de comunicación y del sector de la defensa, estudia cómo luchar contra la manipulación de los usuarios y las noticias falsas, liderando proyectos de detección de noticias falsas.

En el plano del análisis de contenidos, el equipo **LINKMEDIA** investiga la recuperación y el rastreo de imágenes, así como la minería de imágenes, junto con la minería de textos para ayudar a la detección y el rastreo de noticias falsas.

El equipo **ALMANACH** aplica técnicas de procesamiento del lenguaje natural para identificar noticias falsas.

A nivel de gestión de datos y conocimientos, los equipos **CEDAR** y **GRAPHIK** trabajan en el acceso fácil a fuentes de datos heterogéneas para facilitar la verificación y la validación de la información en entornos complejos.

El equipo **DANTE** se centra en el análisis de gráficos para identificar las redes de difusión de noticias falsas y las cuentas que legitiman las noticias falsas.

Los equipos de visualización como **ILDA** estudian los mecanismos necesarios para que los analistas humanos que trabajan de forma colaborativa y dinámica saquen el máximo provecho de los datos, su procesamiento relacionado y los algoritmos de gestión, con el fin de aumentar la seguridad de la información en la industria de los medios de comunicación y otras áreas.

2.4 Modelización de amenazas y ataques con árboles de ataque

[Resumen]

Para representar todo tipo de ataques, se puede utilizar una representación gráfica llamada “árboles de ataque”. En esta representación, cada hoja del árbol expresa un paso que el atacante debe realizar para llevar a cabo su ataque. Cada nodo no terminal contiene una etiqueta que expresa cómo están conectados sus nodos hijos (and, or, sequence). Globalmente, un árbol de ataque representa las secuencias de acciones posibles que el atacante puede realizar para alcanzar su objetivo. Los árboles de ataque se utilizan en gran medida durante la etapa de análisis de riesgos: se identifican los riesgos (amenazas), se diseña una política de seguridad y, a continuación, se eligen los mecanismos de seguridad para implementar dicha política.

En el campo de la seguridad, los llamados "árboles de fallos" se propusieron a principios de los años 80 [44] [VGRH81] para representar gráficamente los riesgos de seguridad a los que se enfrenta un sistema. En 1999, Bruce Schneier adaptó esta representación al campo de la seguridad, introduciendo los "árboles de ataque". [45][Sch99]

Como cualquier árbol, un árbol de ataque está compuesto por nodos y hojas. La etiqueta de un nodo ("AND" u "OR") expresa cómo están conectados los nodos hijo. Cada hoja expresa un paso que el atacante tiene que realizar para llevar a cabo su ataque. Para poder expresar vínculos más sutiles entre las distintas etapas de un ataque, también se han introducido operadores más evolucionados, como el "Sequential AND" que impone un orden temporal entre los operadores de un AND lógico.

Los árboles de ataque comunes aportan una representación cualitativa de un ataque. Algunas investigaciones trabajan para enriquecer los árboles de ataque añadiendo atributos cuantitativos que incluyen el impacto del ataque, los costes de las contramedidas, etc.

La semántica de los árboles de ataque se ha ampliado en gran medida durante la última década. Una de las principales extensiones considera no sólo la descripción de un ataque, sino también las acciones que el administrador de seguridad puede llevar a cabo para detener su progresión. Esta clase de árboles se denominan árboles de ataque-defensa.[46][KMRS10] La originalidad de esta representación radica en que las hojas son heterogéneas y representan tanto la perspectiva del atacante como la del defensor. El vínculo entre las perspectivas del atacante y del defensor también se ha investigado en el ámbito de la correlación de alertas.

Debido a la dificultad de construir árboles de ataque, se han hecho varios intentos de generarlos automáticamente.

Se han propuesto muchas otras técnicas de modelado gráfico, desde pequeñas variaciones de los árboles de ataque originales hasta cambios de representación, como en los gráficos de ataque, los gráficos de ataque bayesianos o las redes de Petri. Para obtener información más detallada sobre los distintos modelos gráficos que se utilizan en seguridad para modelar un ataque y sus sutilezas, el lector debe consultar. [47][KPCS14]

44. [VGRH81] William E. Vesely, Francine F. Goldberg, Norman H. Roberts y David F. Haasl. Fault tree handbook. Technical report, Nuclear Regulatory Commission Washington DC, 1981..
45. [Sch99] Bruce Schneier. Attack Trees: Modeling security threats. Dr. Dobb's Journal, 24(12):21-29, Diciembre 1999.
46. [KMRS10] Barbara Kordy, Sjouke Mauw, Sasa Radomirovic, y Patrick Schweitzer. Foundations of Attack-Defense Trees. In Formal Aspects of Security and Trust, Lecture Notes in Computer Science, páginas 80-95. Springer, Berlin, Heidelberg, Septiembre 2010.
47. [KPCS14] Barbara Kordy, Ludovic Pietre-Cambacedes, y Patrick Schweitzer. DAG-based attack and defense modeling: Don't miss the forest for the attack trees. Computer Science Review, 13(Supplement C):1-38, noviembre 2014.

[Equipos Inria] Modelización de amenazas y ataques con árboles de ataque

↗ El equipo **CIDRE** propone una solución para construir automáticamente reglas de correlación a partir de los árboles de ataque existentes y de la descripción de los sistemas monitorizados.

↗ El equipo **DIVERSE** trabaja en la síntesis de árboles de ataque para apoyar el análisis de riesgo asistido por ordenador.

↗ El equipo **PRIVATICS** utiliza árboles de ataque para la cuantificación de la privacidad.

↗ El equipo **TAMIS** utiliza árboles de ataque para la cuantificación de la seguridad.

Primitivas criptográficas, esquemas y protocolos

La criptografía tiene como objetivo proporcionar técnicas y herramientas para asegurar las comunicaciones, incluso en presencia de un adversario. Históricamente, el principal objetivo de la criptografía era garantizar la confidencialidad de los mensajes mediante el cifrado, es decir, que la información permanezca oculta para las personas no autorizadas. Los primeros métodos de encriptación eran, en general, bastante ingenuos, como el cifrado César, que consiste en cambiar cada letra por una constante (por ejemplo, sustituyendo la "A" por la "D", la "B" por la "E", y así sucesivamente) y puede descifrarse fácilmente utilizando técnicas como el análisis de frecuencias. Un avance sustancial se produjo durante la Segunda Guerra Mundial con la máquina de rotores Enigma utilizada por Alemania. Descifrar el cifrado Enigma requirió un esfuerzo y recursos considerables. Hoy en día, la criptografía se basa en fundamentos matemáticos sólidos y pretende garantizar muchas más propiedades que la mera confidencialidad: la criptografía proporciona herramientas para proteger la integridad y la autenticidad de los mensajes (evitando, por ejemplo, que se modifique el importe de una transacción financiera), para garantizar la irrenunciabilidad (el emisor no puede negar ser el autor de un mensaje) y el anonimato. Además, estas herramientas pueden combinarse para lograr objetivos aún más complejos. El objetivo del criptoanálisis es "romper" las

Primer plano de un conjunto de rotores de Enigma. – TedColes vía Wikipedia, CC0

técnicas criptográficas que están destinadas a garantizar su seguridad, pero en realidad se utiliza para verificar la robustez de éstas.

El capítulo está estructurado en tres partes:

- Las primitivas criptográficas son los bloques de construcción más básicos; dichas primitivas permiten cifrar o firmar digitalmente un mensaje;
- Los esquemas criptográficos generalmente se basan en primitivas para proporcionar objetivos de seguridad más fuertes, garantizando la integridad y autenticidad de mensajes de tamaño arbitrario;

- Los protocolos criptográficos se basan en esquemas para lograr objetivos de seguridad más complejos, por ejemplo, establecer un canal de comunicación seguro que pueda utilizarse para intercambios de mensajes confidenciales y autenticados

3.1 Primitivas criptográficas

[Resumen]
Las primitivas criptográficas, como las funciones de cifrado y las firmas digitales, son los elementos básicos para diseñar sistemas seguros. Pueden dividirse en dos familias: criptografía simétrica y asimétrica. La criptografía simétrica supone que las partes que se comunican comparten en privado una clave secreta. Este tipo de criptografía es más eficiente que la asimétrica, pero requiere que previamente se realice un intercambio de claves seguro. La criptografía asimétrica, o de clave pública, no requiere compartir una clave secreta, ya que la clave pública (utilizada para cifrar o verificar las firmas) no necesita ser secreta. El uso de estos dos tipos de primitivas es complementario. Un procedimiento habitual consiste en intercambiar una clave simétrica utilizando la criptografía asimétrica, con el fin de cifrar las comunicaciones posteriores de forma más eficiente con la criptografía simétrica utilizando la clave intercambiada. Por lo tanto, en la mayoría de las aplicaciones se requieren ambos tipos de criptografía.

Hoy en día disponemos de construcciones consolidadas tanto para la criptografía simétrica como para la asimétrica. Sin embargo, sigue siendo necesario investigar tanto en el criptoanálisis como en el diseño. El objetivo del cripto-analista es encontrar los puntos débiles y evaluar la robustez de las construcciones existentes. Por un lado, este trabajo consiste en evaluar cuidadosamente la dificultad de resolver los problemas subyacentes supuestamente difíciles desde el punto de vista algorítmico, teniendo en cuenta tanto la evolución de la potencia de procesamiento como la mejora de los algoritmos. Por otro lado, este trabajo también debe explorar nuevos métodos de ataque, como los ataques que dependen de un computador cuántico o los ataques que explotan los canales laterales. El diseño de nuevas primitivas puede estar impulsado por avances cripto-analíticos (predecibles) (por ejemplo, debido a la construcción de un computador cuántico) o por una nueva demanda o necesidad de la industria. Entre las nuevas demandas se encuentran los esquemas criptográficos ligeros que puedan funcionar en dispositivos de bajo consumo, así como nuevas funcionalidades, como la necesidad de un cifrado funcional u homomórfico que permita realizar cálculos sobre datos cifrados, lo que suele ser necesario cuando se subcontrata el cálculo de datos sensibles, que es una tendencia actual.

Las primitivas criptográficas son de dos tipos diferentes. Por un lado, las primitivas que se basan en un único secreto compartido por todos los participantes se conocen como criptografía de clave secreta, también llamada criptografía

simétrica. Esta familia también incluye las funciones hash criptográficas: como cualquier función hash, una función hash criptográfica mapea un valor de tamaño arbitrario a uno de tamaño fijo, pero además garantiza propiedades como por ejemplo ser una función unidireccional (es difícil encontrar la pre-imagen de un valor hash) o la resistencia a las colisiones (es difícil encontrar dos valores que mapeen al mismo valor hash). Las funciones hash no se basan en ningún secreto, sino en principios de diseño similares a los del cifrado simétrico. Por otro lado, las primitivas que utilizan claves asimétricas (como la clave de firma y la clave de verificación en primitivas de firma) se reúnen bajo el nombre de criptografía de clave pública, también llamada criptografía asimétrica.

Las primitivas simétricas son muy eficientes tanto en las implementaciones de software como de hardware, por lo que suelen estar bien adaptadas a entornos restringidos. Por el contrario, las primitivas asimétricas requieren herramientas matemáticas complejas para ser implementadas, y su eficiencia es de varios órdenes de magnitud peor que las primitivas criptográficas simétricas. Sin embargo, no requieren un acuerdo previo entre los usuarios. Por lo tanto, las primitivas simétricas y asimétricas ofrecen características de seguridad complementarias, por lo que suelen estar asociadas en aplicaciones concretas para lograr tanto las propiedades de seguridad adecuadas como los requisitos de eficiencia. Por ejemplo, en la mayoría de las aplicaciones, los datos se cifran mediante un esquema de cifrado simétrico eficiente bajo una clave secreta. Esta clave secreta (corta), que debe ser compartida por los usuarios, se transmite con un esquema de cifrado asimétrico o se obtiene mediante un protocolo de intercambio de claves (por ejemplo, el intercambio de claves Diffie-Hellman).

3.1.1 La criptografía en la actualidad

La mayoría de las necesidades básicas de la criptografía se solventan hoy en día con un número relativamente pequeño de primitivas estandarizadas, a saber, AES[48[DR02]] (junto con algún modo de operación como CTR o GCM) para el cifrado simétrico y SHA-2 y SHA-3 para el hash criptográfico. Para la firma digital, el intercambio de claves y el cifrado de clave pública, la mayoría de las primitivas, como RSA[49[RSA78]] o el intercambio de claves Diffie-Hellman[50[DH76]], se basan en supuestos de la teoría de los números (por ejemplo, la dificultad computacional de la factorización o el cálculo de logaritmos discretos).

48[DR02] J. Daemen and V. Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard.* Information Security and Cryptography. Springer, 2002.
49[RSA78] R. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM,* 21(2):120-126, 1978.
50[DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. Information Theory,* 22(6):644-654, 1976.

¿CUÁLES SON LOS PROBLEMAS?

Además de la investigación de nuevas características (por ejemplo, el cifrado funcional), una de las principales tareas de la investigación criptográfica es mantener la confianza en las primitivas existentes e idear otras nuevas siempre que sea necesario. Esta confianza se deriva de la comprensión de las amenazas y su evolución gracias a los nuevos métodos cripto-analíticos o a los avances tecnológicos. Por lo tanto, las primitivas existentes deben ser examinadas continuamente para comprobar que todas estas amenazas se afronten adecuadamente.

El criptoanálisis nos dice cuándo y por qué las primitivas deben evolucionar (por ejemplo, con claves más grandes, más rounds) o ser reemplazadas (por ejemplo, debido a los avances algorítmicos o tecnológicos). Proporciona al diseñador herramientas de evaluación y criterios de diseño que ayudan a adaptar las primitivas criptográficas existentes o a proponer otras nuevas que sean resistentes a los nuevos ataques.

3.1.2 Criptoanálisis

El objetivo del criptoanálisis académico es comprender las amenazas a la seguridad de las primitivas existentes para adelantarse a los adversarios maliciosos. Una de las dificultades es que las amenazas pueden evolucionar con el tiempo, con el progreso de los algoritmos, las matemáticas o las computadoras (por ejemplo, la ley de Moore o la computación cuántica), pero las capacidades del atacante también evolucionan (por ejemplo, el acceso físico a una implementación, que no siempre se tuvo en cuenta, ahora debe tenerse en cuenta en el caso de las primitivas ligeras para la IoT). El criptoanálisis es la base de la confianza que tenemos en estas primitivas: cuanto más las analizamos, más podemos confiar en ellas. Proporciona una medida empírica de seguridad gracias a un escrutinio exhaustivo e incesante, en busca de posibles debilidades. El conocimiento del criptoanálisis de vanguardia es, pues, la columna vertebral para el diseño de primitivas seguras. A continuación, distinguiremos entre el criptoanálisis matemático, que se centra en el diseño, y el criptoanálisis de implementación, que explota detalles particulares de la implementación. También distinguiremos entre el criptoanálisis clásico y el cuántico.

CRIPTOANÁLISIS CLÁSICO

El criptoanálisis de una primitiva criptográfica consiste en resolver un problema que se considera difícil en el modelo de seguridad (por ejemplo, recuperar la clave secreta, descifrar o falsificar una firma sin la clave, encontrar una colisión). El ataque tiene éxito si es más eficiente de lo esperado según los requisitos de seguridad. El criptoanálisis normalmente sólo consigue romper versiones reducidas de la primitiva (por ejemplo, una clave más pequeña, un tamaño de bloque menor, menos rounds). El margen de seguridad de una determinada primitiva

se cuantifica en función de cuánto difieren las versiones reducidas de la primitiva original. El criptoanálisis de vanguardia es en realidad el único criterio de seguridad para decidir en qué momento una primitiva debe evolucionar o ser sustituida. Por ejemplo, después de 10 años de esfuerzo cripto-analítico por parte de varios equipos de investigación, se ha encontrado recientemente una colisión para la función hash SHA-1[51], que ahora se considera insegura desde el punto de vista criptográfico y está siendo sustituida lentamente en las aplicaciones. Otro reciente y espectacular avance cripto-analítico, que afecta en gran medida a la seguridad de algunos cripto-sistemas basados en el emparejamiento, es un algoritmo cuasi-polinómico para calcular logaritmos discretos en campos finitos de pequeñas características .

Dado que la criptografía moderna aboga por la seguridad por conocimiento, es decir, por tener un razonamiento público del diseño, incluyendo la justificación de las elecciones de diseño, el criptoanálisis público es la única manera de poder confiar en tales primitivas. Por lo tanto, cualquier intento de estandarización va precedido de una larga e intensa fase de criptoanálisis que es necesaria para evaluar la futura norma. Además, el análisis de las nuevas propuestas suele conducir a la definición de nuevos ataques,que también podrían considerarse nuevas amenazas para las primitivas existentes.

[Secciones Destacadas] Records en cálculos de criptoanálisis

PRIVATICS Inria ha sido un actor clave en varios cómputos pioneros para el criptoanálisis:

↗ En 2010 se logró factorizar una clave RSA de 768 bits. El esfuerzo computacional fue de unos 1500 años de CPU y 2 años naturales. Este cómputo sin precedentes fue dirigido por la EPFL y el equipo **CARAMBA** de Inria. Más o menos al mismo tiempo, las claves de las tarjetas de crédito pasaron de 896 a 960 bits y la ANSSI publicó una recomendación para utilizar claves RSA de 2048 bits en 2010.

↗ En 2013, los investigadores de los equipos **CARAMBA** y **OURAGAN** en colaboración con A. Joux diseñaron un algoritmo eficiente para romper logaritmos discretos en campos finitos de la característica 2, así como emparejamientos para curvas algebraicas basadas en campos binarios finitos.

↗ Desde 2014, los investigadores de **CARAMBA** y **GRACE** han resuelto logaritmos discretos en varios campos finitos del GF(pk) para k≤6.

↗ En 2015, el ataque FREAK (equipo **PROSECCO**) pone de manifiesto cómo la factorización de claves RSA de 512 bits (en combinación con un error de implementación común) puede utilizarse para romper las conexiones TLS afectando alrededor del 25% de la web. Este trabajo dio lugar a correcciones en los principales navegadores y sitios web.

↗ En 2015, el ataque LogJam a TLS (equipos **PROSECCO** y **CARAMBA**) pone de manifiesto

51. https://www.inria.fr/centre/saclay/actualites/sha-1-les-predictions-d-inria-verifiees

que, utilizando un cálculo previo para un grupo específico de 512 bits, las claves Diffie Hellman pueden romperse de forma efectiva y que las claves (aún muy utilizadas) en grupos de 768 bits están ahora al alcance de los equipos de investigación académica.
↗ En 2016, el ataque SLOTH (equipos **PROSECCO** y **SECRET**) pone de manifiesto cómo las colisiones de hash en MD5 y SHA-1 pueden utilizarse para romper la autenticación basada en firmas en protocolos como TLS. Este trabajo condujo a la eliminación de MD5 y SHA-1 en TLS 1.3.
↗ En 2016, los investigadores del equipo **CARAMBA**, en colaboración con colegas de la Universidad de Pensilvania, mostraron cómo resolver Diffie-Hellman para números primos especialmente elaborados. Como resultado, unos meses más tarde, se retiró una RFC (Request For Comments) que no proporcionaba detalles sobre la generación de los parámetros.

CRIPTOANÁLISIS CUÁNTICO

Después de un par de décadas de estudio mayoritariamente limitado al mundo académico, la computación cuántica está ahora en el centro de una carrera entre empresas de alta tecnología como Google, Microsoft o IBM. Aunque la perspectiva de una gran máquina cuántica universal está todavía a muchos años de distancia, descartar su potencial impacto en la criptografía se ha convertido en una posición bastante insostenible. De hecho, el Instituto Nacional de Estándares y Tecnología de Estados Unidos, NIST, ha lanzado recientemente un llamado a la creación de cripto-sistemas resistentes a los ataques realizados con la ayuda de una computadora cuántica.

La sabiduría común en este campo dice que la criptografía de clave pública que depende de la dificultad de la factorización o del logaritmo discreto, por ejemplo, RSA o curvas elípticas, falla irremediablemente en un mundo cuántico debido al algoritmo de Shor[52[Sho]]. La criptografía simétrica parece en principio mucho más inmune al criptoanálisis cuántico, ya que la principal aceleración aplicable parece venir dada por el algoritmo de Grover en la búsqueda exhaustiva, una tarea en la que los computadores cuánticos sólo proporcionan una ventaja cuadrática, es decir, el costo de una búsqueda exhaustiva baja de N a √N donde N es el número de claves. En particular, se podría pensar ingenuamente que duplicar el tamaño de la clave secreta es suficiente para hacer frente a los ataques cuánticos contra la criptografía simétrica. Desgraciadamente, esta apreciación no está respaldada por los muchos años de trabajo necesarios para ganar confianza en la seguridad de los cripto-sistemas mediante un criptoanálisis específico. Recientemente han aparecido nuevos resultados en este sentido, por ejemplo, mostrando que, en ciertos modelos, el algoritmo cuántico de Simon para la búsqueda de periodos rompe completamente la seguridad de los modos de funcionamiento más

52[Sho] P. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *IEEE FOCS'94*.

utilizados (como CBC-MAC o GCM) para la autenticación y el cifrado autenticado.

La lección es que el criptoanálisis cuántico merece mucha más atención de la que ha recibido hasta ahora. Los estándares tardan muchos años en evolucionar debido a las nuevas amenazas. Por lo tanto, ahora es necesario comprender mejor el impacto de los computadores cuánticos en la criptografía para poder determinar cómo se deben modelar los ataques cuánticos, para entender cuán seguras son las soluciones postcuánticas basadas en retículas o códigos, y de manera más general, para integrar las técnicas cuánticas en la caja de herramientas de los diseñadores de cripto-sistemas.

CRIPTOANÁLISIS DE LA IMPLEMENTACIÓN

La seguridad de una primitiva criptográfica depende tanto del software como del hardware que implementa la primitiva y del hardware en el que se ejecuta el código. El atacante de las primitivas criptográficas puede limitarse a tener acceso al software, por ejemplo, ejecutando remotamente otro proceso en la misma máquina, o también puede tener acceso al hardware y ser capaz de medir alguna cantidad física durante la ejecución, como el tiempo exacto o el voltaje o incluso realizar la inyección de fallos. Sin embargo, la distinción entre software y hardware no es tan significante para los criptógrafos, que establecen un modelo de las amenazas que suele abstraerse de los detalles de bajo nivel y de las diferencias entre el software y el hardware. En la sección 2.1 se ofrece un análisis más orientado a la práctica de este tipo de ataques.

Cuando un algoritmo criptográfico se implementa en un dispositivo físico (una tarjeta inteligente, un teléfono móvil, un computador, etc.), un adversario puede medir las propiedades físicas del sistema durante la ejecución del algoritmo criptográfico. La precisión de estas medidas depende de si el atacante tiene acceso físico al hardware. En este caso, son capaces de medir las variaciones del consumo de energía o de las radiaciones electromagnéticas durante la ejecución. Incluso, pueden ser capaces de realizar inyecciones de fallos. Si el atacante sólo tiene acceso al software, por ejemplo, mediante la ejecución remota de algún otro proceso en el mismo dispositivo, puede obtener sólo cierta información de temporización. Sin embargo, en ambos casos, estas medidas físicas están correlacionadas con la clave secreta manipulada por el sistema y la clave puede ser extraída eficientemente a menos que se utilicen contramedidas especiales. Por lo tanto, un importante campo de investigación se centra en la evaluación de la eficacia de estos ataques y en la elaboración de contramedidas eficaces. En particular, ahora conocemos bien la protección que ofrece el enmascaramiento, que utiliza técnicas de cálculos multipartitos para dividir el secreto en partes que no están correlacionadas con el secreto real.

Los ataques físicos también incluyen los ataques de fallo, en los que un atacante

manipula el circuito para inducir un error (por ejemplo, manipulando la fuente de alimentación o con un disparo láser) y explota la diferencia entre una salida normal y una defectuosa para recuperar la clave secreta. La mayoría de los ataques de implementación requieren acceso físico al sistema, pero también son posibles los ataques cuando un adversario puede simplemente ejecutar código en la misma máquina que la víctima. En particular, los ataques a la caché y a los fallos basados en la técnica Rowhammer [53][KDK+14] pueden ser muy efectivos y se han demostrado a partir de código JavaScript ejecutado en un navegador web.

3.1.3 Diseño

Las nuevas primitivas son diseñadas o bien cuando se produce algún avance criptoanalítico (por ejemplo, la ruptura de la mayoría de las funciones hash estandarizadas en 2004 y 2005) o bien para responder a alguna demanda urgente de la industria (por ejemplo, primitivas ligeras para dispositivos de bajo costo). En la criptografía moderna, las nuevas primitivas vienen acompañadas de argumentos de diseño y seguridad. Estos argumentos no proporcionan una garantía incondicional de seguridad y pueden adoptar formas diferentes. En la criptografía asimétrica, el argumento suele determinar que cualquier adversario que rompa la primitiva con unos parámetros determinados resolverá un problema que se considera ampliamente difícil (por ejemplo, la factorización de un número de 2048 bits correspondiente al producto de dos primos, la descodificación genérica de un código lineal o la detección de un vector corto en una retícula de parámetros determinados). En la criptografía simétrica, los argumentos se basan en las propiedades de los bloques de construcción subyacentes que garantizan (o tienden a garantizar) su resistencia a clases conocidas de ataques (por ejemplo, ataques lineales, diferenciales o algebraicos).

[Secciones Destacadas] Concursos criptográficos internacionales

Durante más de 20 años, los nuevos estándares criptográficos principales se han establecido tras concursos abiertos iniciados por organismos de normalización o proyectos internacionales. Estos concursos atraen propuestas de muchos países, tanto del mundo académico como de la industria. Los candidatos son entonces examinados durante varios años por toda la comunidad criptográfica en un proceso público de evaluación de la seguridad. Inria ha presentado varias primitivas a estos concursos y ha contribuido al proceso de evaluación.

↗ El equipo **SECRET** presentó dos cifrados de flujo al proyecto eSTREAM, puesto

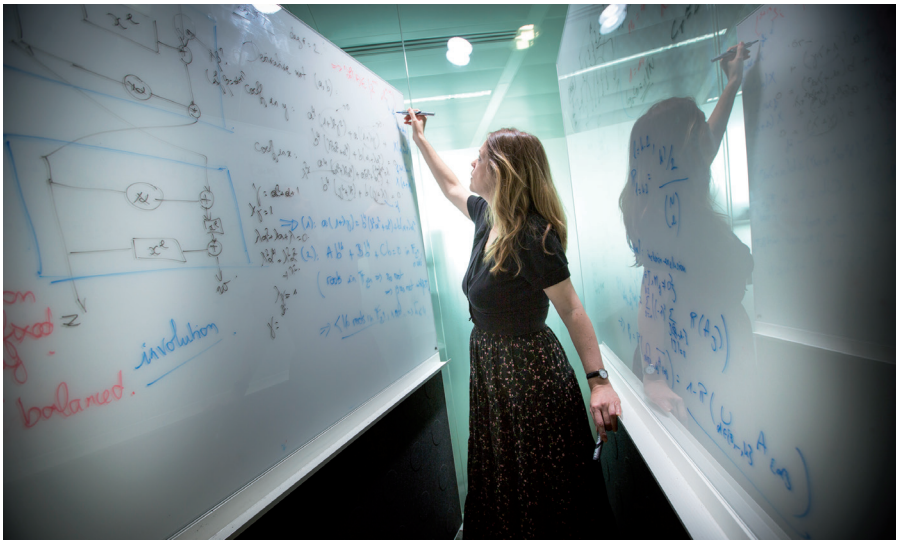
53[KDK+14] Y. Kim, R. Daly, J. Kim, Ch. Fallin, J.-H. Lee, D. Lee, Ch. Wilkerson, K. Lai, and O. Mutlu. Flipping bits in memory without accessing them: An experimental study of DRAM disturbance errors. In *ACM/IEEE ISCA'14,* 2014.

en marcha en 2004 por la Red de Excelencia ECRYPT. Uno de ellos, Sosemanuk, ha sido seleccionado (entre 34 propuestas) en la cartera final de cifrados recomendados para entornos orientados al software[a].

↗ Las funciones hash Shabal y SIMD, diseñadas por los equipos **SECRET** y **CASCADE** junto con varios socios, son dos de las 64 candidatas al concurso SHA-3[b] lanzado por el NIST en 2007. Ambas fueron seleccionadas entre las 11 semifinalistas del concurso.

↗ Los equipos **ARIC, GRACE, POLSYS** y **SECRET** participan en el diseño de 10 (de 68) candidatos al proceso de estandarización de la criptografía postcuántica[c] iniciado por el NIST en 2017.

a. http://www.ecrypt.eu.org/stream/
b. https://csrc.nist.gov/projects/hash-functions/sha-3-project
c. https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization



Propiedades criptográficas de los componentes de un sistema de cifrado – © Inria / Foto C. Morel

CRIPTOGRAFÍA LIGERA

En las últimas décadas, hemos asistido a un enorme aumento del número de dispositivos inteligentes de bajo costo: por ejemplo, tarjetas sin contacto, llaveros (RKE), redes de sensores, domótica, etiquetas NFC/RFID o implantes médicos. La mayoría de ellos transmiten información sensible de forma inalámbrica (por ejemplo, las tarjetas sin contacto suelen utilizarse para el pago, el control de acceso o el cobro de tasas) y requieren criptografía para hacerlo de forma segura. Por

desgracia, las restricciones de hardware de los dispositivos embebidos limitan su capacidad de cálculo y su energía disponible (al estar alimentados por una batería o incluso de forma pasiva), lo que impide el uso de la criptografía convencional. Por ello, muchos productos industriales utilizan criptografía casera ligera (por ejemplo, MIFARE Classic, KeeLoq, Megamos, Hitag2) o ninguna criptografía (por ejemplo, implantes médicos, ratones inalámbricos). Para llenar este vacío, se han diseñado cifrados ligeros y seguros que funcionan con un bajo consumo de recursos críticos (energía, potencia, tiempo de ejecución). Estas estrictas limitaciones de implementación pueden tener inconvenientes, como un tamaño de bloque más pequeño que obligue a renovar las claves más a menudo o una baja latencia. En las últimas décadas se han introducido y normalizado varios diseños, como KASUMI (UMTS), PRESENT (ISO/IEC 29192-2) o HIGHT (ISO/IEC 18033-3). La criptografía simétrica ligera es un área de investigación muy activa, con varias propuestas nuevas cada año y un intento de normalización en curso por parte del NIST[54]. También es necesario un esfuerzo importante para crear diseños ligeros de primitivas de intercambio de claves y de criptografía asimétricas. Una de las especificidades de la criptografía ligera es que, en la mayoría de las aplicaciones, los dispositivos de bajo coste son muy vulnerables a los ataques físicos. Por lo tanto, las primitivas ligeras no sólo deben tener una especificación ligera, sino también una implementación segura.

En un entorno tan restringido, la generación de aleatoriedad también es un problema. Aunque las monedas aleatorias son necesarias en la mayoría de los esquemas criptográficos como también para asegurar su implementación de hardware, son bastante difíciles de generar, especialmente a bajo costo. Los generadores pseudo-aleatorios permiten expandir una semilla aleatoria en un flujo mayor de bits aleatorios, pero hay que estudiar tanto la entropía de la semilla como la calidad de la expansión. Cualquier debilidad que se produzca puede debilitar todo el sistema. Es incluso peor si el adversario puede tener algún control sobre el dispositivo, lo que podría permitirle reducir la entropía del estado interno.

CRIPTOGRAFÍA POSTCUÁNTICA

La disponibilidad de la computación cuántica dejará obsoletas todas las primitivas criptográficas basadas en la teoría de los números que se utilizan hoy en día de forma rutinaria y casi exclusiva para asegurar las comunicaciones. Aunque los computadores cuánticos tardarán una o varias décadas en ser una realidad, la comunidad investigadora tiene que ponerse en marcha ya y empezar a crear primitivas alternativas, especialmente para los mecanismos de intercambio de

54. https://www.nist.gov/programs-projects/lightweight-cryptography

claves y las firmas digitales. En particular, las primitivas que las reemplazarán deben estar listas pronto si se desea asegurar la confidencialidad a largo plazo[55]. El NIST ha puesto en marcha una iniciativa[56] para estandarizar los esquemas criptográficos resistentes a la computación cuántica –postcuánticos– (cifrado de clave pública, intercambio de claves, firma digital). Se espera que la fase de análisis público, de cinco años de duración, en el marco del proceso de normalización del NIST proporcione una mejor visión sobre el nivel de seguridad y el rendimiento de estas técnicas.

La técnica más antigua de seguridad cuántica para la criptografía asimétrica, el esquema de cifrado McEliece, se propuso en 1978 y es contemporáneo de la RSA. Su seguridad está relacionada con la dificultad de la decodificación de un código lineal arbitrario y es el trabajo fundamental de la criptografía basada en códigos. Posteriormente, se desarrolló la criptografía multivariada, basada en la dificultad de la resolución de sistemas polinómicos, y la criptografía basada en retículos, que se basa en la dificultad de encontrar vectores cortos en un retículo euclidiano. La última década ha sido extremadamente productiva, en particular con la aparición de LWE [57][Reg] y sus variantes cíclicas mucho más prácticas (Ring-LWE). Estas técnicas están alcanzando la madurez y se impondrán como alternativa práctica para la criptografía asimétrica en la próxima década. Para ser exhaustivos, mencionemos que se están considerando otras técnicas para la criptografía postcuántica, siendo una de las más notables la criptografía basada en hash (basada en la técnica del árbol de Merkle) que permite esquemas de firma digital resistentes a la computación cuántica siempre que la función hash criptográfica sobre la que se construye sea resistente a las colisiones.

[Desafío de investigación 3] Criptografía postcuántica

Se cree que la construcción de un computador cuántico universal (es decir, no de propósito especial) será factible en las próximas décadas. Por tanto, es importante pensar ahora en la criptografía resistente a la computación cuántica, ya que cierta información que se cifra hoy puede seguir siendo sensible dentro de, por ejemplo, 50 años. La mayor parte de la criptografía asimétrica que se utiliza hoy en día se basa en la dificultad de la factorización o en el cálculo de logaritmos discretos, problemas que se sabe que un computador cuántico puede resolver con facilidad. Por lo tanto, es necesario buscar alternativas: las primitivas basadas en retículos, en códigos y en

55.Obsérvese que la situación es crítica para la confidencialidad, ya que un adversario puede almacenar hoy en día documentos encriptados que necesitan ser protegidos durante décadas, incluso después de que se disponga de una computadora cuántica.
La situación es diferente en el caso de la integridad, que puede reforzarse mediante la renuncia a documentos una vez que la amenaza cuántica se haga realidad.

56. https://csrc.nist.gov/Projects/Post-Quantum-Cryptography

57 [Reg] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. En Proceedings of the 37th ACM Symposium on Theory of Computing (STOC 2005).

multivariantes son las candidatas más destacadas. Es urgente realizar un análisis de seguridad en profundidad de estos nuevos esquemas.

CRIPTOGRAFÍA BASADA EN LAS LEYES DE LA FÍSICA

La mayoría de las propuestas criptográficas se basan en supuestos computacionales y, por tanto, son vulnerables a los avances algorítmicos y de hardware. Una forma de dejar de depender por completo de esas suposiciones computacionales es utilizar la criptografía cuántica. Utilizando la comunicación cuántica, es posible construir un protocolo de distribución de claves incondicionalmente seguro (conocido como el protocolo BB84). Esto significa que incluso un adversario todopoderoso (potencialmente cuántico) no puede romper el esquema. Esto ofrece una seguridad a muy largo plazo, pero sólo puede utilizarse para un número limitado de aplicaciones, como la distribución de claves, debido a las limitaciones de despliegue, por lo que suele combinarse con la criptografía estándar o postcuántica. En Europa y en Asia se desarrollan redes cuánticas para poder realizar protocolos de distribución de claves cuánticas incondicionales. Actualmente, estos protocolos sólo funcionan en una distancia limitada de unos 50-150 km. Para crear redes a gran escala, se necesitan nodos de confianza, que pueden ser peligrosos desde el punto de vista criptográfico, o repetidores cuánticos. Los repetidores cuánticos están tecnológicamente fuera de alcance hoy en día, pero parecen mucho más fáciles de construir que una computadora cuántica completa y podrían ser una realidad en un futuro cercano.

Otra solución para obtener criptografía incondicionalmente segura es utilizar otras leyes de física. La criptografía relativista [58][Kan15], por ejemplo, utiliza restricciones espacio-temporales entre los agentes para realizar protocolos criptográficos con seguridad incondicional. Estas restricciones limitan las posibles aplicaciones, pero, a diferencia de la criptografía cuántica, pueden desarrollarse a gran escala con la tecnología actual. Es importante seguir los avances de la criptografía incondicional, que puede verse como un plan alternativo seguro en caso de que perdamos la confianza en los supuestos computacionales utilizados en la criptografía estándar.

58 [Kan15] J. Kaniewski. Relativistic quantum cryptography. Tesis doctoral, Centro de Tecnologías Cuánticas, Universidad de Singapur, 2015. https://arxiv.org/pdf/1512.00602.pdf.

3.2 Esquemas criptográficos

[Resumen] Esquemas criptográficos

Mientras que las primitivas criptográficas son las bases de construcción básicas, los esquemas criptográficos logran propiedades más robustas con modos de operación específicos. Algunas aplicaciones, como la computación externalizada, también pueden requerir funcionalidades más avanzadas que el cifrado clásico. Los llamados esquemas de cifrado homomórfico y funcional permiten trabajar con datos encriptados, y las pruebas criptográficas ("pruebas de conocimiento") pueden utilizarse para obtener pruebas de que el cálculo externalizado se ha realizado correctamente. Con la creciente complejidad de los esquemas criptográficos y sus pruebas de seguridad, ha surgido una nueva tendencia, denominada criptografía asistida por computadora (computer-aided cryptography), que consiste en desarrollar herramientas para comprobar las pruebas de seguridad y lograr una mayor confianza en la seguridad de algunas construcciones.

La seguridad de un cifrado por bloques, como AES, o de una función trampa unidireccional, como la función RSA, no suele proporcionar esquemas de cifrado seguros por sí solos. Por ejemplo, estas funciones básicas no son aleatorias y cifrar dos veces lo mismo dará lugar a dos textos cifrados idénticos, con lo que se filtrará la información de que los dos textos planos eran idénticos. Por lo tanto, los esquemas de cifrado o los esquemas de firma suelen estar definidos por una primitiva junto con un modo de funcionamiento que especifica cómo utilizar la primitiva para acomodar mensajes de longitud arbitraria y alcanzar un objetivo de seguridad específico. Algunos ejemplos de estas construcciones son los modos de cifrado en bloque estandarizados para el cifrado simétrico (por ejemplo, CBC, CTR) o para el cifrado autenticado (por ejemplo, CCM, GCM), algunos esquemas de relleno para el cifrado asimétrico (por ejemplo, OAEP[59][FOPS01]) o para las firmas (por ejemplo, PSS). El nivel de seguridad que ofrecen estas construcciones se evalúa bajo el supuesto de que la primitiva subyacente tiene un comportamiento ideal. El objetivo es garantizar que un esquema dado es seguro siempre que no se haya identificado ninguna debilidad específica para la primitiva. Hay dos enfoques complementarios para analizar la seguridad de un modo de operación: la búsqueda de ataques genéricos (es decir, independientes de la primitiva subyacente) proporciona límites superiores en el nivel de seguridad, mientras que las pruebas de seguridad proporcionan límites inferiores.

59 [FOPS01] E. Fujisaki, T. Okamoto, D. Pointcheval y J. Stern. RSA-OAEP is secure under the RSA assumption. En Proceedings of Crypto '01, volumen 2139 de LNCS, 2001.

3.2.1 Construcciones demostrables

El área de la criptografía demostrable, iniciada en el trabajo pionero de Goldwasser y Micali [60][GM84], tiene sus raíces en la teoría de la complejidad computacional. Los adversarios se modelan como máquinas de Turing de tiempo polinomial arbitrario, que tienen acceso a oracles (para cifrar o descifrar). La seguridad se expresa típicamente como la incapacidad del adversario para distinguir (con una probabilidad significativamente mejor que 1/2) si tiene acceso a un oracle de cifrado o a una función que siempre devuelve una cadena aleatoria. Las pruebas se realizan por reducción: un adversario que puede ganar un juego de indistinguibilidad puede utilizarse para construir eficientemente un adversario que pueda invertir la función unidireccional subyacente o distinguir el cifrado de bloques subyacente de una permutación aleatoria. Por lo tanto, la prueba muestra esencialmente que romper la construcción es tan difícil como resolver el problema subyacente que supuestamente es difícil. Tales resultados se han obtenido para los modos de operación clásicos utilizados para lograr autenticación simétrica y cifrado asimétrico. Sin embargo, todavía existe una importante línea de investigación sobre el diseño de nuevos modos de operación eficientes para el cifrado simétrico con un alto nivel de seguridad. En efecto, la mayoría de los modos de operación de cifrado por bloques tienen su seguridad limitada por el llamado birthday-bound: se vuelven inseguros si el número de llamadas al cifrado por bloque subyacente se acerca a 2n/2, donde n es el tamaño del bloque. Esto es un problema importante para la criptografía ligera, ya que en estas aplicaciones se prefieren los cifrados de bloque que funcionan con bloques de 64 bits. Esto implica que la cantidad de datos que se pueden cifrar con la misma clave debe ser muy inferior a 32 GBytes. Este problema se ha demostrado recientemente en el modo de funcionamiento CBC utilizado en HTTP sobre TLS y OpenVPN en un ataque denominado Sweet32[61] que llevó al NIST a bajar el límite antes de cifrar de nuevo 3DES a 8 MBytes. El diseño de un modo de operación eficiente con un nivel de seguridad más alto es, por tanto, un importante problema que queda por resolver. Aunque este enfoque se ha utilizado ampliamente para demostrar la seguridad de los esquemas de cifrado de clave pública y de los esquemas de firma bajo supuestos bien conocidos, ahora es necesario para cualquier nueva construcción, y en concreto para las construcciones avanzadas como el cifrado totalmente homomórfico y el cifrado funcional.

60 [GM84] S. Goldwasser y S. Micali. Probabilistic encryption. Journal of Computer and System Sciences, 28(2), 1984.

61. https://sweet32.info/

3.2.2 Cifrado homomórfico y funcional

Con el desarrollo del almacenamiento y la computación externalizados, el cifrado clásico no es suficiente: cuando son cifrados, se garantiza la privacidad de los datos, pero no se puede realizar ningún tipo de procesamiento con ellos. En 2009, Gentry [62][Gen09] propuso el primer esquema de cifrado que permite realizar operaciones homomórficas: a partir del cifrado de dos mensajes, es posible producir el cifrado de la suma o del producto, sin ninguna información secreta. Más concretamente, es posible enviar datos encriptados a la nube y dejar que ésta evalúe un circuito sobre estos datos encriptados. La nube puede entonces devolver el cifrado del resultado, sin haber obtenido ninguna información sobre el mismo. El propietario de la clave de descifrado puede entonces descifrar el resultado. Mientras que la construcción inicial del cifrado totalmente homomórfico era prohibitiva tanto a nivel computacional como comunicacional, recientemente se han propuesto muchas mejoras, y ahora se pueden evaluar concretamente algunos circuitos pequeños.

Sin embargo, el resultado obtenido por la nube sigue estando encriptado bajo la misma clave que las entradas, por lo que sólo puede ser compartido con aquellos que ya pueden desencriptar las entradas. Esta es la razón por la que Boneh, Sahai y Waters [63][BSW11] propusieron la noción de cifrado funcional: una autoridad puede distribuir claves de desencriptado funcional que ayudan a calcular la evaluación de una función dada en el texto plano. Sin embargo, las claves no revelan la entrada completa, sino sólo el resultado de la función calculada. La autoridad cifra inicialmente los datos con una clave maestra y obtiene una clave funcional kf para una función elegida f. La función f se evalúa sobre los datos del texto plano cuando se descifra con kf. Esto permite, por ejemplo, hacer alguna agregación de datos (análisis estadístico) sin revelar los datos.

Mientras que las técnicas clásicas (logaritmo discreto) permiten instaurar el cifrado funcional para familias de funciones sencillas, como el producto interior [64][ABDP15], parece que se necesitarán técnicas basadas en retículas para funciones avanzadas. Del mismo modo, el cifrado ElGamal es (simplemente) homomórfico, pero para lograr un cifrado totalmente homomórfico, parece necesario el cifrado basado en retículas, o algunos enfoques con ruido/error. De ahí la intensa investigación sobre la criptografía basada en retículas y en códigos.

62 [Gen09] C. Gentry. Fully homomorphic encryption using ideal lattices. In Proceedings of the 41st ACM Symposium on Theory of Computing (STOC'09), 2009. http://portal.acm.org/citation.cfm? id=1536414.1536440
63 [BSW11] D. Boneh, A. Sahai, y B. Waters. Functional encryption: Definitions and challenges. En *Proceedings of Theory of Cryptography Conference (TCC'11),* 2011. http://ia.cr/2010/543.
64 [ABDP15] M. Abdalla, F. Bourse, A. De Caro, y D. Pointcheval. Simple functional encryption schemes for inner products. *En Proceedings of Public Key Cryptography (PKC'15), 2015.*

[Desafío de investigación 4] Computación sobre datos encriptados

La necesidad de computar sobre datos encriptados ha surgido, en particular, con la aparición de la nube y la computación externalizada. En criptografía, es
te problema puede resolverse mediante un cifrado homomórfico o funcional. En 2009, Gentry demostró en su innovador artículo que era posible construir un esquema de cifrado totalmente homomórfico (FHE). Sin embargo, esta construcción seguía siendo teórica y resultaba completamente impracticable debido a su escaso rendimiento. Desde entonces, se han hecho progresos significativos en los esquemas FHE, logrando aproximadamente una velocidad, aún muy baja, de 50 puertas lógicas por segundo. Los avances significativos tendrán aplicaciones extremadamente útiles para la computación en la nube que preserva la privacidad, por otro lado, cualquier avance técnico puede ser rápidamente explotado como una ventaja económica.

3.2.3 Pruebas de conocimiento

Para constatar la exactitud de los cálculos externalizados realizados por la nube en la que no confiamos, se necesitan pruebas para convencer al usuario del comportamiento honesto de la nube. Ha habido una gran actividad en las pruebas de zero-knowledge o witness-indistinguishable por cuestiones de privacidad. Sin embargo, en el contexto de la computación externalizada, la solidez y la precisión de la prueba son la cuestión más importante, ya que no hay preocupaciones de privacidad con respecto al usuario, de ahí la nueva primitiva SNARG (argumento sucinto no interactivo) que proporciona pruebas sucintas para declaraciones complejas.

No obstante, las pruebas habituales de zero-knowledge son de gran interés para las credenciales anónimas y cualquier tipo de mecanismo de autenticación avanzada que preserve el anonimato. Recientemente se han desarrollado nuevas técnicas con las Funciones Smooth-Projective Hash (SPHF).

Además, para limitar las interacciones, las Pruebas de Non-Interactive Zero-Knowledge (NIZK) se han vuelto más eficaces, primero con la metodología de Groth-Sahai, que permite probar muchos tipos de relaciones entre escalares o elementos de grupos comprometidos, o más recientemente con las NIZK Casi-Adaptativas, basadas en SPHF que son más específicas pero más compactas y eficientes.

3.2.4 Criptografía asistida por computadora

La realización de pruebas de reducción de forma rigurosa es extremadamente difícil porque manipulan complejos algoritmos probabilísticos. En efecto, la literatura contiene muchos ejemplos de errores sutiles en las pruebas, siendo un ejemplo famoso la construcción OAEP, "probada" en 1994 con un error descubierto

en 2001. En su documento, Halevi [65][Hal05] aboga por el uso de asistentes de pruebas para verificar la corrección de las partes triviales de las pruebas criptográficas automáticamente. En cuanto a las reducciones en la teoría de la complejidad, las pruebas constan de una parte creativa para encontrar la reducción y una parte trivial, pero difícil, que consiste en verificar la corrección de la reducción. En los últimos años, varias herramientas, como CryptoVerif[66], CertiCrypt y EasyCrypt[67], han demostrado que estas partes son realmente susceptibles de ser verificadas de forma automática. La más madura y versátil de estas herramientas es EasyCrypt: la herramienta consiste en un probador de teoremas interactivo específico que muestra propiedades relacionales en esquemas modelados como programas probabilísticos. Aunque la herramienta ofrece sólidas garantías, también requiere un alto nivel de conocimiento. Para aplicaciones concretas, como la seguridad de los métodos de cifrado contra un texto sencillo seleccionado o un cibertexto, se ha logrado una automatización completa con la herramienta específica ZooCrypt, lo que ha dado lugar a la verificación de numerosos mecanismos y al diseño de otros nuevos que han demostrado ser seguros. El área de la criptografía asistida por computadora se está expandiendo actualmente gracias a la ampliación de su espectro, incluyendo la seguridad de las primitivas contra los canales laterales, las construcciones basadas en emparejamientos, la aplicación de estas ideas a los protocolos criptográficos en lugar de a los métodos (véase el apartado 4.3) y la mejora de la automatización.

[Equipos Iria] Primitivas y esquemas criptográficos

↗ El equipo **ARIC** está trabajando en la criptografía basada en retículas (LBC). La algoritmia reticular es un área de investigación consolidada que está siendo revitalizada por LBC y por las nuevas herramientas y conceptos que introdujo. Su objetivo es contribuir al gran cambio tecnológico que supone pasar de la criptografía convencional a la basada en retículas.

↗El equipo **CARAMBA** estudia los aspectos matemáticos, algorítmicos y de software de alto rendimiento de la criptografía asimétrica basada en la teoría de los números (criptosistemas RSA y Diffie-Hellman, curvas elípticas). Sus trabajos de criptoanálisis demuestran la urgencia de aumentar el tamaño de las claves de varias de estas primitivas. El equipo también participa en el diseño y el criptoanálisis de primitivas criptográficas simétricas (en particular en el contexto ligero).

↗El equipo **CASCADE** se centra en los aspectos de seguridad demostrable de las primitivas avanzadas o en entornos avanzados. En concreto, estudian las primitivas que preservan la privacidad (como la FHE, el cifrado funcional, etc.), pero también tienen en cuenta

65 [Hal05] Sh. Halevi. A plausible approach to computer-aided cryptographic proofs. Technical Report 181, IACR Cryptology ePrint Archive, 2005.
66 http://cryptoverif.inria.fr
67 https://www.easycrypt.info/

a los adversarios poderosos, con ataques de canal lateral y computadoras cuánticas.
↗ El equipo **GRACE** trabaja en la teoría algorítmica de los números y en las cuestiones computacionales relacionadas con las curvas algebraicas sobre diversos campos y anillos aritméticos. También construyen códigos para la corrección de errores. Su objetivo es proporcionar mejores criptosistemas y mejores evaluaciones de seguridad para sus tamaños de clave.
↗ El equipo **LFANT** investiga algoritmos en teoría de números y geometría aritmética. Abarcan todos los aspectos, desde la teoría de la complejidad, pasando por las implementaciones optimizadas, hasta las aplicaciones criptológicas.
↗ El objetivo del equipo **MARELLE** es estudiar y utilizar técnicas de verificación de pruebas matemáticas en la computadora para garantizar la corrección del software. Han aplicado sus técnicas en el contexto de las pruebas criptográficas contribuyendo al desarrollo del asistente de pruebas de propósito especial EasyCrypt.
↗ El equipo **OURAGAN** trabaja en cálculos efectivos de objetos algebraicos con aplicaciones a la criptología.
↗El equipo **POLSYS** desarrolla algoritmos eficientes para calcular las soluciones complejas o reales en campos finitos. La criptología es una de las muchas aplicaciones, donde se puede utilizar en el tema emergente del criptoanálisis algebraico. Este consiste en reducir la seguridad de un criptosistema a la resolución de un sistema algebraico con coeficientes en un campo finito
↗ El equipo **SECRET** trabaja en el diseño y análisis de primitivas simétricas, de primitivas asimétricas basadas en códigos de corrección de errores y en esquemas criptográficos basados en las leyes de la física. Han contribuido al diseño de varias primitivas (cifrados de flujo, cifrados por bloques, funciones hash, criptografía basada en código y esquemas de cifrado, e intercambio de claves), y también a muchos trabajos de criptoanálisis en estas áreas. Se centran especialmente en el diseño de primitivas seguras desde el punto de vista cuántico e investigan el uso de algoritmos cuánticos para atacar esquemas tanto simétricos como asimétricos.

3.3 Protocolos y servicios criptográficos: hacia una seguridad demostrable

[Resumen]

Hoy en día, la seguridad de las comunicaciones y transacciones está garantizada por protocolos criptográficos, como TLS. Sin embargo, la seguridad de las primitivas y esquemas criptográficos subyacentes no es suficiente para garantizar los objetivos generales de seguridad, como la confidencialidad, la autenticidad o el anonimato. Ni siquiera un examen minucioso de estos protocolos por parte de expertos puede garantizar la ausencia de vulnerabilidades: por lo tanto, las

pruebas de seguridad rigurosas, posiblemente asistidas por computador, desde especificaciones hasta implementaciones, se han vuelto indispensables para garantizar aún más el nivel de confidencialidad. Podemos distinguir tres enfoques en este ámbito. El primero utiliza pruebas por reducción para demostrar que romper la seguridad de un protocolo implicaría romper las primitivas criptográficas subyacentes. Se trata de pruebas matemáticas, generalmente escritas a mano, aunque una nueva tendencia consiste en utilizar técnicas de demostración de teoremas y de verificación de programas para conseguir pruebas comprobadas por computador. La segunda línea de investigación utiliza herramientas de verificación automatizada para analizar las especificaciones de los protocolos y encontrar vulnerabilidades en la lógica del protocolo, como los ataques Man-in-the-Middle. Estas herramientas son capaces de analizar protocolos complejos, pero idealizan la criptografía subyacente. Por último, el tercer enfoque tiene como objetivo producir implementaciones verificadas. Este enfoque se basa en sistemas de tipos expresivos para lenguajes de programación específicos y requiere una gran experiencia, pero puede dar lugar a una implementación verificada de comienzo a fin. Un éxito importante en este ámbito es la implementación completamente verificada de TLS.

Una criptografía sólida no es en sí misma suficiente para garantizar los objetivos de seguridad a un nivel superior, por ejemplo, asegurar las comunicaciones o las transacciones web. Utilizar y programar correctamente con criptografía es una tarea complicada y hay muchos ejemplos de vulnerabilidades de seguridad que no requieren romper la criptografía subyacente (véanse los ejemplos de ataques Heartbleed, French Passport y TLS más abajo). Por lo tanto, es importante diseñar y analizar los estándares de los protocolos y las bibliotecas que hacen uso de la criptografía para garantizar propiedades de alto nivel.

Los protocolos criptográficos, como TLS (Transport Layer Security), IKE (Internet Key Exchange) o Kerberos, se encargan de asegurar nuestras conexiones y transacciones web. Son programas distribuidos que utilizan la criptografía para garantizar, por ejemplo, la confidencialidad de los datos transmitidos y la autenticación de las comunicaciones y las entidades. Con la creciente diversidad de servicios electrónicos, éstos se están extendiendo rápidamente: por ejemplo, son la base de la seguridad de las aplicaciones de mensajería y de los objetos con RFID, como los pasaportes electrónicos; también son fundamentales en servicios de seguridad como el ampliamente utilizado Single-Sign-On (SSO) o los servicios basados en la nube.

Asegurar el intercambio de datos en Internet – © Inria / Foto C. Morel

El diseño de protocolos y normas de seguridad requiere conocimientos en varias áreas de la informática, como la criptografía, las redes informáticas y también la implementación segura. Esta tarea es difícil, incluso para los expertos, que pueden pasar por alto ataques debido a la gran complejidad de estos protocolos. Una de las dificultades para diseñar e implementar correctamente los protocolos criptográficos proviene del hecho de que la seguridad debe garantizarse en presencia de un atacante arbitrario que controle la red y pueda comprometer a los participantes del protocolo. Las vulnerabilidades pueden surgir en todos los niveles. Por ejemplo, el famoso ataque Heartbleed[68] se debe a un error de implementación, que permite un desbordamiento de memoria, en la popular implementación OpenSSL de TLS: este ataque no reveló un error en la especificación del protocolo, ni rompió la criptografía subyacente. Se demostró que una de las primeras versiones del pasaporte electrónico francés[69] era vulnerable a un linkability attack, que permitió rastrear a los titulares de los pasaportes. La vulnerabilidad se debió a imprecisiones en la especificación del protocolo en lo que respecta a los mensajes de error: el pasaporte electrónico francés utilizaba

68. https://en.wikipedia.org/wiki/Heartbleed
69. Defects in e-passports allow real-time tracking. The Register, 26 de enero de 2010. http://www.theregister.co.uk/2010/01/26/epassport_rfid_weakness/

mensajes de error detallados, lo que permitía diferenciar un pasaporte concreto, previamente observado, de otro, mediante un ataque de repetición. Por último, los ataques FREAK[70] y LogJam[71] contra TLS mezclan vulnerabilidades a diferentes niveles para minimizar las longitudes de clave criptográfica. Tratar correctamente el código heredado para garantizar la compatibilidad con versiones anteriores, pero evitar los downgrade attacks es extremadamente complicado.

Como indican los ejemplos anteriores, diseñar correctamente protocolos y servicios y garantizar su seguridad es una tarea difícil. Por lo tanto, se necesitan pruebas de seguridad rigurosas y técnicas de análisis formal para mejorar su seguridad. Ha habido diferentes enfoques complementarios y líneas de investigación: algunos analizan los protocolos a nivel de especificación, mientras que otros analizan directamente la implementación del protocolo; el análisis puede centrarse en la lógica del protocolo o en la criptografía subyacente; el grado de automatización también puede variar, desde pruebas completamente escritas a mano, hasta análisis totalmente automatizados, así como pruebas interactivas comprobadas por computador.

3.3.1 Seguridad demostrable para protocolos criptográficos

El enfoque de la seguridad demostrable, introducido inicialmente para dar garantías sólidas de seguridad a los esquemas criptográficos, se basa en pruebas reduccionistas (inspiradas en las pruebas de complejidad computacional): normalmente, se demuestra que romper la primitiva criptográfica es al menos tan difícil como romper un problema subyacente computacionalmente difícil, como la factorización, el cálculo de logaritmos discretos, etc. Este enfoque se ha aplicado a los protocolos, donde la reducción muestra, por ejemplo, que romper una propiedad de seguridad esperada es tan difícil como romper una primitiva criptográfica subyacente. Los adversarios que se consideran son máquinas de Turing probabilísticas arbitrarias de tiempo polinómico que pueden interactuar de distintas maneras con los participantes legítimos del protocolo y comprometer a algunos de ellos.

Asimismo, se han propuesto marcos generales para el diseño riguroso de protocolos de seguridad. Los protocolos de computación multipartita segura (MPC) proporcionan un modelo general para calcular el resultado de una función, permitiendo que distintas partes que no confían entre sí proporcionen una entrada confidencial. Los protocolos MPC permiten implementar una gran variedad de protocolos, pero las construcciones eficientes, seguras contra adversarios fuertes, siguen siendo un tema de investigación activo y desafiante. Otros frameworks, llamados universalmente componibles, o basados en la simulación, tienen como

70. https://en.wikipedia.org/wiki/FREAK
71. https://en.wikipedia.org/wiki/Logjam_(computer_security)

objetivo ser altamente modulares y mostrar la seguridad de los componentes que pueden ser ensamblados en sistemas más grandes. Ser capaz de dividir la prueba de un sistema complejo en pruebas de sus componentes puede allanar el camino para crear protocolos seguros por diseño que puedan utilizarse como bloques de construcción para sistemas más grandes.

En el trabajo descrito anteriormente, las pruebas se realizan generalmente a mano, lo que las hace propensas a los fallos. En cuanto a los esquemas criptográficos, ha habido iniciativas para automatizar las pruebas en estos modelos, por ejemplo, a través de la herramienta CryptoVerif, o para utilizar pruebas teóricas interactivas específicas, por ejemplo, EasyCrypt. Mejorar el alcance y la automatización de estas herramientas sigue siendo un campo de investigación activo y desafiante.

3.3.2 Análisis simbólico automatizado de las especificaciones de los protocolos criptográficos

El análisis simbólico automatizado de los protocolos criptográficos se centra en la lógica del protocolo y su comportamiento concurrente y puede aplicarse a especificaciones complejas de protocolos. Aunque los objetivos son similares a los descritos en la sección anterior, las técnicas y los modelos subyacentes difieren. En los modelos simbólicos se supone que el llamado atacante Dolev-Yao [72][DY81] controla la red de comunicación por completo: un atacante puede leer cualquier mensaje enviado en la red, eliminar mensajes e insertar (o modificar) mensajes. El atacante es computacionalmente ilimitado, pero las primitivas criptográficas son idealizadas: la forma en que un atacante puede manipular los mensajes está explícitamente dada por un conjunto de reglas. Dichas reglas suelen especificar que cuando el atacante conoce un cifrado y la correspondiente clave de descifrado puede extraer el texto plano, pero no se permite otra operación no especificada (como el criptoanálisis). Por lo tanto, estos modelos manipulan las primitivas criptográficas a un nivel abstracto y axiomático. Sin embargo, los llamados resultados de solidez [73][AR07] vinculan este enfoque con el de la seguridad demostrable (mostrando que una prueba simbólica implica la existencia de un cálculo), aunque su alcance es bastante limitado. Apoyándose en las técnicas del razonamiento automatizado y de la teoría de la concurrencia, las pruebas en modelos simbólicos pueden, a menudo, ser completamente automatizadas, explorando todos los posibles comportamientos de los atacantes. Hoy en día existen herramientas automatizadas maduras, como ProVerif[74], así como Tamarin[75]

72 [DY81] D. Dolev y A. Yao. On the security of public key protocols (extended abstract). En Proceedings of FOCS'81, 1981.
73 [AR07] Martin Abadi y Phillip Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). J. Cryptology, 20(3), 2007.
74. http://proverif.inria.fr
75. https://tamarin-prover.github.io/

y AVISPA[76], para analizar muchos objetivos de seguridad. Las principales áreas de investigación en este campo consisten en aplicar estas herramientas a una clase más amplia de propiedades de seguridad (técnicamente esto requiere mostrar equivalencias de programas de comportamiento) que nos permitan analizar el anonimato y las propiedades de privacidad, considerar ejecuciones de protocolos en plataformas parcialmente comprometidas y escalar a protocolos con un complejo flujo de control subyacente.

3.3.3 Implementaciones de protocolos verificados

Como ilustra el ataque Heartbleed, los errores de implementación pueden introducir graves vulnerabilidades. Por ello, se han utilizado técnicas del campo de los lenguajes de programación para estudiar directamente la seguridad de las implementaciones. Este enfoque se basa principalmente en el uso de sistemas de tipos expresivos para establecer teoremas de seguridad directamente en el nivel de implementación. Un enfoque particular que ha tenido éxito en esta área se basa en el lenguaje F*[77], un lenguaje efectivo fuertemente tipado de orden superior, especialmente diseñado para desarrollar implementaciones probadas. Por ejemplo, el lenguaje F* se ha utilizado en el proyecto miTLS[78], una colaboración entre Microsoft Research e Inria que ha dado como resultado una implementación de referencia completamente probada del protocolo TLS 1.2 y del nuevo protocolo candidato para TLS 1.3, llegando hasta el nivel de la implementación de las primitivas criptográficas. Actualmente, este enfoque sólo se aplica a implementaciones cuidadosamente escritas con muchos tipos de anotaciones. Un reto importante en este ámbito es hacerlo aplicable a un código más general, el cual podría no haber sido escrito por un experto en métodos formales. Otra línea de investigación consiste en extraer modelos directamente de las implementaciones, ya sea especificando pequeños subconjuntos del lenguaje o, en un entorno más específico, probando el sistema en modo "caja negra". Estos modelos se pueden entonces analizar automáticamente.

3.3.4 Voto electrónico por Internet

Terminamos este capítulo con una aplicación particular de los protocolos criptográficos: el voto electrónico. Las elecciones son, sin duda, la piedra angular de las democracias modernas y un proceso crítico para la seguridad. Estonia ha sido pionera en esta práctica desde 2005, y utiliza las elecciones por Internet incluso para las elecciones al parlamento nacional. Algunas regiones de Suiza y Australia también ofrecieron el uso del voto por Internet. En Francia, se propuso

76. http://www.avispa-project.org/
77. https://www.fstar-lang.org/
78. https://mitls.org/

el voto por internet a los ciudadanos franceses que viven en el extranjero durante las elecciones nacionales de 2012, pero en las elecciones de 2017 no se renovó esta opción, por motivos de seguridad.

Las principales garantías de seguridad que deben ofrecer unas elecciones son el secreto del voto y la exactitud del resultado. El voto secreto debe garantizar que nadie sepa cómo ha votado un determinado votante (a menos que pueda deducirse del resultado de la elección, por ejemplo, en caso de voto unánime). La corrección garantiza que el resultado se corresponde con el recuento de los votos, tal y como los expresaron todos los votantes con derecho a voto. En las elecciones tradicionales que utilizan papeletas, al menos en Francia, estas propiedades se garantizan a través de un ritual de votación, con una urna transparente, una cabina de votación que proporciona la privacidad necesaria para emitir un voto en secreto, y observadores que controlan la urna y el recuento. El uso de computadoras y máquinas complica significativamente esta tarea, ya que los programas informáticos pueden contener errores y la corrección de un sistema es difícil de asegurar. Además, el software puede ser manipulado intencionadamente, o un malware puede alterar su funcionalidad. Esto puede cambiar el resultado de los votos, así como filtrar los votos individuales de las personas, rompiendo así las dos propiedades fundamentales de unas elecciones.

Para superar los problemas mencionados, se han propuesto elecciones secretas y verificables de extremo a extremo, reforzadas criptográficamente. El secreto se consigue generalmente emitiendo un voto encriptado. Este voto se mezcla con otros votos antes del descifrado, de modo que ya no puede vincularse a la identidad del votante, o el recuento se realiza de forma homomórfica, es decir, el recuento se calcula sobre los votos cifrados proporcionando un cifrado del resultado, sin un descifrado de los votos individuales. Esto garantiza que ni el escrutador ni el servidor que recopila los votos pueden romper el secreto del voto. La corrección se consigue mediante la noción de elecciones verificables de extremo a extremo: el votante puede verificar que su voto se ha registrado correctamente y que el recuento se ha realizado correctamente. Para ello, el sistema genera pruebas criptográficas de que las operaciones se han realizado correctamente. Esta propiedad evita tener que verificar la corrección del software que realiza el recuento, ya que genera pruebas, es decir, pruebas matemáticas que pueden ser verificadas independientemente de la corrección del resultado.

[Secciones Destacadas] El sistema de voto electrónico Belenios

El equipo **PESTO** ha trabajado en la definición precisa de las propiedades que debe garantizar un sistema de votación y en la verificación formal de estas. Este trabajo permitió descubrir un ataque contra el popular sistema de votación Helios y también aclaró los supuestos de confianza de muchos protocolos. Los equipos **PESTO** y **CARAMBA** desarrollaron el sistema de votación Belenios: un sistema de votación gratuito y de

código abierto que garantiza el secreto y la verificabilidad de los votos, incluida la verificabilidad de la elegibilidad de los votos emitidos, evitando así la manipulación de las papeletas. Sin embargo, este sistema, como muchos otros, tiene defectos: no impide la coerción, ya que un votante puede demostrar cómo ha votado. Por lo tanto, el sistema sólo se recomienda para elecciones de baja coerción. Además, es vulnerable al malware que pueda instalarse en la máquina que ejecuta el cliente votante: un malware de este tipo podría filtrar el voto, rompiendo el secreto, o cambiar el voto antes de que sea encriptado (y la verificabilidad sólo permite rastrear la papeleta encriptada). Resolver la resistencia a la coerción y la resistencia al malware de forma satisfactoria sigue siendo una cuestión de investigación abierta.

Asimismo, hay que tener en cuenta que el voto por Internet elimina la garantía de privacidad de una cabina de votación y requiere un medio para identificar a los votantes a distancia. En realidad, este es el caso de cualquier sistema de voto a distancia, incluidos los basados en papeletas, como el voto por correo. Por último, los sistemas criptográficos de voto electrónico también requieren que los votantes confíen en algunos expertos, ya que se basan en nociones matemáticas avanzadas, lo que hace que su comprensión sea difícil. Por lo tanto, estos sistemas no parecen estar aún preparados para elecciones de alto nivel, por ejemplo, políticas.

[Desafío de investigación 5] Protocolos criptográficos extremo a extremo verificados formalmente

Dado que la seguridad de los protocolos criptográficos es extremadamente difícil de garantizar (las pruebas de lápiz y papel suelen contener errores), el uso de métodos rigurosos y formales se perfila cada vez más como la única forma de alcanzar el nivel de seguridad esperado para esta clase de sistemas. Por lo tanto, el ámbito de las pruebas de seguridad asistidas por computador es un tema cada vez más importante y debe incluir todos los aspectos, desde la especificación hasta la implementación. Los trabajos recientes, en particular en torno a TLS 1.3, han demostrado que esto ya es posible.

Sin embargo, estas pruebas siguen requiriendo un código cuidadosamente elaborado y un nivel muy alto de conocimientos. Aprovechar las técnicas de demostración para hacerlas aplicables a un código más general y utilizables por un público más amplio es ahora el principal reto. Los distintos protocolos suelen garantizar diferentes propiedades de seguridad, como el anonimato, no tienen todavía la misma madurez que las herramientas para verificar las propiedades de autenticación. Otro reto es considerar modelos de adversarios más fuertes, por ejemplo, un adversario que pueda controlar parte de la computadora a través de un *malware*.

[Equipos Inria]

Hay varios equipos de Inria que trabajan en pruebas formales y análisis de protocolos criptográficos:

↗ El equipo **CASCADE**, además de su trabajo sobre primitivas criptográficas, también trabaja en el diseño de protocolos y su análisis de seguridad en modelos computacionales.

↗ El objetivo del equipo **PESTO** es construir modelos y técnicas formales para el análisis y el diseño asistidos por computadora de protocolos de seguridad, utilizando técnicas de razonamiento automatizado, teoría de la concurrencia y lenguajes de programación. Están especialmente interesados en el análisis automatizado de las propiedades de anonimato y los protocolos de voto electrónico. Contribuyen al desarrollo de varias herramientas, como AVISPAa, DEEPSECb y el prover Tamarinc.

↗ El equipo **PROSECCO** lleva a cabo investigaciones de seguridad formales y prácticas sobre protocolos criptográficos, seguridad de software, seguridad web y mecanismos de protección de hardware. Para ello, diseñan e implementan lenguajes de programación, herramientas de verificación formal, monitores dinámicos, marcos de pruebas, compiladores verificados, etc. Desarrollan el prover automatizado de protocolos ProVerifd y contribuyen al diseño del lenguaje F* que utilizaron para desarrollar una implementación completamente verificada de TLS. También han trabajado en la extracción de modelos de implementaciones del estándar de gestión de claves PKCS#11, lo que ha dado lugar a la creación de la startup Cryptosense.

a. http://www.avispa-project.org/
b. https://deepsec-prover.github.io/
c. https://tamarin-prover.github.io/
d. http://proverif.inria.fr

4

Servicios y mecanismos de seguridad

Cuando un usuario empieza a utilizar un sistema informático, primero se identifica (identificación) y luego demuestra que realmente es quien dice ser (autenticación). A continuación, el sistema utiliza esta identidad comprobada para conceder o restringir el acceso a un recurso o servicio sólo a los usuarios o entidades autorizados (control de acceso), o para evitar que una información específica llegue a un destino determinado (control de flujo) mediante las acciones de un usuario. Cuando un usuario necesita ejecutar un programa en una máquina que no es de confianza, se puede utilizar un mecanismo de hardware para garantizar el aislamiento y la integridad del software. Todos estos servicios de seguridad son preventivos. Desgraciadamente, a veces pueden ser burlados por los atacantes, por lo que también se necesita una seguridad reactiva. Así, las acciones de los usuarios se supervisan para comprobar que no violan la política de seguridad (detección de intrusos y correlación de alertas). Por supuesto, dicha violación de la política puede haberse producido sin el conocimiento del usuario, que puede haber sido atacado por un malware que actúe en su nombre. Por lo tanto, el análisis y la detección de malware son también servicios de seguridad que deben ofrecerse. Lo ideal es que, si se detecta una intrusión o un malware, el sistema debería reaccionar, al menos para reconfigurarse y evitar otro ataque similar.

En el resto de este capítulo, expondremos sucesivamente los servicios de seguridad presentados anteriormente: identificación y autenticación (apartado 4.1), control de acceso y de flujo (apartado 4.2), computación confiable (apartado 4.3), detección de intrusos y correlación de alertas (apartado 4.4), análisis y detección de malware (apartado 4.5) y reacción (apartado 4.6).

4.1 Identificación y autenticación

[Resumen]

La identificación y la autenticación son generalmente los dos primeros servicios de seguridad utilizados al iniciar un intercambio cibernético, ya sea entre un ser humano y una máquina o entre dos máquinas.

La identificación, para una entidad determinada (es decir, un usuario, un servicio, un dispositivo, etc.), es el acto de declarar su identidad. La autenticación, para esta entidad, es el acto de demostrar que es realmente la entidad que previamente ha afirmado ser.

La autenticación se utiliza para restringir el acceso a un recurso o servicio sólo a los usuarios o entidades autorizadas. La autenticación se consigue presentando un autenticador que generalmente pertenece a una de las tres clases siguientes:

lo que se sabe, por ejemplo, una contraseña o un pin; lo que se tiene, por ejemplo, una tarjeta de acceso; lo que se es, por ejemplo, mecanismos basados en la biometría. Estos autenticadores suelen combinarse en la llamada autenticación multifactorial. A pesar de sus numerosos inconvenientes, las contraseñas siguen siendo el medio de autenticación más común.

Otra forma de identificación es la propiedad de un dato. Esto puede lograrse mediante la marca de agua, que consiste en ocultar mensajes en los datos.

Una buena técnica de marca de agua debe crear un vínculo robusto entre el soporte y el mensaje oculto, de manera que la distorsión del soporte no borre el mensaje. La marca de agua se utiliza en la gestión de los derechos de autor y la protección anti-copia, centrándose principalmente en los contenidos multimedia, aunque el espectro de aplicaciones es más amplio y puede utilizarse, por ejemplo, para proteger los códigos fuente, o las bases de datos.

La identificación, para una entidad determinada (es decir, un usuario, servicio, dispositivo, etc.), es el acto de declarar su identidad. Por ejemplo, un usuario debe proporcionar sus credenciales de acceso. Como tal, la identificación no es realmente un servicio de seguridad, ya que la entidad puede mentir y dar información errónea. Por eso es necesario el servicio de autenticación; es el acto de demostrar que uno realmente es quien dice ser. Por ejemplo, un usuario podría revelar la contraseña que está asociada a sus credenciales de acceso.

Obsérvese que algunos autores consideran que la autenticación es un servicio relacionado con la identificación. Como tal, la palabra identificación puede usarse para "autenticación de identidad".

4.1.1 Autenticación de usuarios

La autenticación se utiliza para restringir el acceso a un recurso o servicio sólo a los usuarios o entidades autorizadas. Por ejemplo, para acceder a un sistema se requiere un nombre de usuario y una contraseña[79]. Muchos servicios en línea prestados por empresas privadas o administraciones públicas también requieren de autenticación de usuario. En estos casos, la identificación se realiza a través de una red y los medios de identificación pueden ir unidos a un protocolo criptográfico como TLS (Transport Layer Security).

La autenticación se consigue presentando un autenticador que generalmente pertenece a una de las tres clases siguientes:

- lo que sabe, por ejemplo, una contraseña o un pin;

79. La autenticación de los datos es diferente de la autenticación de entidades y generalmente se logra utilizando medios criptográficos, por ejemplo, mediante un protocolo de intercambio de claves, y el uso de un código de autenticación de mensajes (MAC, por el inglés Message Authentication Code) con la clave resultante, véase el capítulo 3.

- lo que tiene, por ejemplo, una tarjeta de acceso;
- lo que es, por ejemplo, mecanismos basados en biometría.

Estos autenticadores suelen combinarse en la llamada autenticación multifactorial. Por ejemplo, la retirada de efectivo con una tarjeta de crédito requiere tanto la posesión de la tarjeta de crédito como el conocimiento del pin. Crear una nueva identidad de usuario, dejar que el usuario la gestione, ofrecer un servicio de recuperación de contraseñas y, tal vez, la autenticación multifactor es, sin embargo, una tarea muy compleja.

Las contraseñas siguen siendo el medio más común de autenticación. Desgraciadamente, la gestión de contraseñas es complicada, tanto para el usuario que quiere ser identificado como para el sistema que concede el acceso. Por parte del sistema, las contraseñas deben almacenarse. Sin embargo, las contraseñas no deben almacenarse como texto plano[80], ya que una filtración de la lista almacenada comprometería directamente todas las cuentas y contraseñas de los usuarios. Por lo tanto, se recomienda encarecidamente almacenar sólo el valor hash de las contraseñas, utilizando una función hash unidireccional. Además, como las contraseñas deben ser fácilmente recordadas por los humanos, son vulnerables a los ataques de diccionario (o de adivinación), es decir, a los ataques de fuerza bruta por enumeración. Distinguimos entre ataques de adivinación en línea y fuera de línea. En los ataques de adivinación en línea, un atacante prueba todas las contraseñas posibles ejecutando el mecanismo de identificación para cada prueba. Estos ataques pueden frustrarse añadiendo un retraso adicional después de un intento fallido o limitando el número de intentos, lo que hace que este enfoque sea inviable. Los ataques de adivinación fuera de línea son posibles cuando se filtra una lista de hash de contraseñas. Para evitar que el mismo ataque de fuerza bruta se aplique a todas las contraseñas al mismo tiempo, se utiliza un valor de sal: para cada usuario, se almacena un valor aleatorio, llamado sal, y el valor hash del resultado de la concatenación de la contraseña con la sal. Algunas funciones hash también están especialmente diseñadas para ser costosas en tiempo y consumo de memoria, en particular en hardware dedicado, para frenar los ataques de fuerza bruta.

Lo ideal es tener una función hash que sea rápida al verificar contraseñas correctas y lenta en las incorrectas: esta idea ha sido parcialmente realizada por la noción de pimienta. Como en el caso de la sal, se hace un hash de un valor aleatorio adicional, la pimienta. Sin embargo, este valor no se almacena y debe ser encontrado por fuerza bruta: cuando se proporciona la contraseña correcta, el número esperado de hashes es N/2, donde N es el número de posibles valores de pimienta, mientras que se necesitan N hashes para descartar una contraseña incorrecta. Desde el punto de vista del usuario, las contraseñas son a menudo

80 Véase, por ejemplo, Recommendation R22 in ANSSI Note technique - Recommendations pour la securisation des sites web. [en francés] https://www.ssi.gouv.fr/uploads/IMG/pdf/NP_Securite_Web_NoteTech.pdf

complicadas de gestionar: lo ideal es que sean difíciles de adivinar y que no se reutilicen para diferentes servicios. Algunos servicios exigen el uso de dígitos o caracteres especiales en las contraseñas. Sin embargo, recientes recomendaciones[81] cuestionan esta práctica y recomiendan contraseñas más largas, también llamadas frases de contraseña. De hecho, estudios sobre datos reales han demostrado que, sin restricciones, la palabra "password" aparece como una de las opciones más populares. Si se añaden caracteres especiales o dígitos, "password" es sustituido por "¡password!" y "password123" en la lista de contraseñas más frecuentes, por lo que no mejora la seguridad. Una buena práctica es, por tanto, utilizar un gestor de contraseñas que cifre todas las contraseñas utilizando una contraseña maestra. Los ataques no técnicos, como la ingeniería social o el phishing, se analizan en el apartado 2.3.1.

Dado el elevado número de robos de contraseñas, se ha recurrido a eliminarlas por completo o a combinarlas con un segundo autenticador. El uso de la autenticación de dos factores está más extendido. Para pagos en línea, el protocolo 3D-secure puede depender de un código de confirmación, enviado por SMS al teléfono móvil, el código debe volver a introducirse en el dispositivo utilizado para el pago. El objetivo es demostrar tanto el conocimiento del número de la tarjeta de crédito como la posesión del teléfono. Del mismo modo, el protocolo Google 2 Step, puede enviar un código de confirmación, requerir pulsar el teléfono, o ser configurado para solicitar el uso del dispositivo de autenticación FIDO U2F USB conectado a su computadora. Estos protocolos multifactoriales suelen ofrecer más seguridad. Sin embargo, la mayor complejidad tanto del protocolo como de los métodos de recuperación, por ejemplo, cuando se pierde o se rompe un teléfono, puede aumentar a veces la superficie de ataque.

La biometría es otro medio de identificación de usuarios que se basa en sensores para medir características biológicas, como las huellas dactilares, el iris, la voz o las características faciales. Una diferencia inherente con otros autenticadores es que los datos biológicos no son a priori secretos y no pueden ser modificados o sustituidos. Por ello, los sensores son cada vez más sofisticados para distinguir la presencia física de una copia (por ejemplo, un dedo falso con una copia de la huella dactilar o una foto de una cara). Aunque estos medios son cada vez más populares, por ejemplo, las computadoras portátiles y los celulares pueden desbloquearse mediante el reconocimiento de la huella dactilar o, más recientemente, de la cara, siguen padeciendo de falsos negativos y falsos positivos. Por lo tanto, se recomienda que sólo se utilicen como segundo factor de autenticación.

81. NIST Special Publication 800-63B. Digital Identity Guidelines – Authentication and Lifecycle Management – https://pages.nist.gov/800-63-3/sp800-63b.html

4.1.2 Identificación del propietario de los datos: marca de agua

La ocultación de datos es el arte de ocultar mensajes en un medio de cobertura. Abarca dos ramas, la esteganografía y marcas de agua, donde la palabra "ocultación" tiene dos significados diferentes.

En la esteganografía, "ocultar" los mensajes significa que un adversario, el esteganalista, no puede detectar estadísticamente si una porción de datos contiene un mensaje secreto. Así, cambiar o no el valor de un píxel de una imagen compartida por dos o más personas es una forma sencilla (y realmente ingenua) de transmitir información oculta entre estas personas. La esteganografía tiene aplicaciones en los servicios de inteligencia con el estegoanálisis en la lucha contra el terrorismo, por ejemplo.

En las marcas de agua, "ocultar" significa que los mensajes secretos están profundamente incrustados en los datos. Este secreto puede ser invisible para un ser humano, por lo que la marca de agua es un caso especial de esteganografía. Crea un vínculo sólido entre un medio y el secreto, de modo que la distorsión del medio (por ejemplo, girar o recortar una imagen) no borra el secreto. Las marcas de agua se aplican a la gestión de los derechos de autor y a la protección anti-copia. Por ejemplo, una marca de agua identifica al propietario de un dato mediante la incrustación en sus datos de su identidad en forma de mensaje oculto.

Mientras que la literatura sobre marcas de agua se centra principalmente en el contenido multimedia, el espectro de las naturalezas de los medios de cobertura es muy amplio: programas (protección de códigos fuente, parámetros de clasificadores de Deep Learning, etc.), bases de datos, mapas, objetos 3D o secuencias de ADN.

Las marcas de agua no deben estropear el uso habitual del medio de cobertura. En el caso de los contenidos multimedia, el usuario no debe percibir ninguna diferencia. La incrustación de marcas de agua hace un gran uso de los modelos perceptivos humanos. En el caso de los programas, la disminución del rendimiento (relevancia del resultado, tiempo de ejecución) debe ser mínima. En el caso de las bases de datos, las consultas deben dar respuestas similares.

La robustez de una marca de agua se determina midiendo cómo aumenta la probabilidad de un error de decodificación del mensaje oculto a medida que el medio protegido se distorsiona más y más. El peor caso de ataque posible se define como el ataque que maximiza esta probabilidad de error para un nivel determinado de distorsión.

La incrustación de marcas de agua no es más que un esquema de comunicación secreta. La incrustación y la descodificación comparten una clave secreta que define cómo el mensaje que se quiere ocultar modula una parte determinada del medio de cobertura. La seguridad entra en escena cuando la misma clave secreta se utiliza para proteger muchos soportes. Se trata de responder a la pregunta de si un adversario puede estimar la clave secreta analizando estos medios con marca de agua. Una vez revelada la clave secreta, el adversario puede borrar la marca de agua (eliminando

cualquier prueba de propiedad), modificar la marca de agua o incrustar su propia marca de agua en cualquier medio para usurpar la propiedad.

La tendencia actual del uso de marcas de agua se movió desde la autenticación de la propiedad a la identificación del consumidor de medios. Esto afecta especialmente a los documentos confidenciales y de gran valor. Gracias a las marcas de agua, el código de identificación de un usuario se incrusta en las copias de su medio para hacerlo único. De este modo, se revela la identidad del usuario que ha filtrado el contenido. Esto no impide la redistribución ilegal en sí, pero es un arma disuasoria para evitar filtraciones. Los códigos de identificación están diseñados de tal manera que incluso si una colusión de varios usuarios mezcla sus copias, la decodificación identificará al menos a uno de los traidores.

Está surgiendo otra tendencia que relaciona la ocultación de datos con la generación de muestras adversas que engañan a los clasificadores de aprendizaje profundo.

Las marcas de agua robustas son una tecnología madura. Muchos resultados de investigación ya se han transferido a productos de la vida real. El número de trabajos de investigación sobre el tema ha disminuido drásticamente en los últimos años. El principal consumidor de la tecnología de marcas de agua es la industria del entretenimiento. Sin embargo, todavía no existe una marca de agua robusta y segura. Parece que la industria del entretenimiento tiene poco interés en la seguridad y está satisfecha con el nivel de robustez alcanzado hasta ahora. Por tanto, no se recomienda seguir investigando sobre marcas de agua como medio para autenticar la propiedad. La tendencia se centra en un diseño multicapa de marcas de agua y rastreo de código de traidores, y en el desarrollo de protocolos entre las partes (propietarios de contenidos, distribuidores de contenidos y consumidores de contenidos), quienes no confían entre sí.

[Equipos Inria] Identificación y autenticación

↗ El equipo **LINKMEDIA** y su empresa derivada Lamark trabajan en la protección multimedia (audio/vídeo) mediante marcas de agua. Por ejemplo, en el ámbito del rastreo de traidores, el equipo propone códigos de identificación tales que, incluso si una colusión de varios usuarios mezcla sus copias, la decodificación identificará al menos a uno de los traidores. Desde el punto de vista del atacante, el equipo estudia cómo un atacante que tiene uno o varios datos con marca de agua puede estimar la clave secreta utilizada por el esquema de marca de agua.

↗ El equipo **MULTISPEECH** estudia la autenticación basada en la voz y la detección de ataques de spoofing. Coorganiza el desafío internacional ASVspoof.

↗ El equipo **PESTO** aplica métodos de verificación formal y simbólica para analizar la seguridad de los protocolos de autenticación multifactor.

↗ El equipo **PROSECCO** utiliza la verificación simbólica en ProVerif para analizar protocolos de autenticación web como OAuth 2.0 y ACME con respecto a un novedoso modelo de amenaza de atacantes web.

4.2 Control de acceso y control de flujo

[Resumen]
La aplicación de la seguridad implica, en primer lugar, definir con precisión qué entidad previamente identificada y autenticada puede tener acceso a qué información y de qué manera. Tradicionalmente, los permisos de acceso (lectura o escritura) a la información se conceden si y sólo si se cumple alguna condición (por ejemplo, el usuario que solicita el acceso está correctamente autenticado). Una secuencia de lectura y escritura genera un flujo de información que también puede ser controlado en algunos casos.

La seguridad preventiva requiere definir con precisión qué entidad puede acceder a qué información y de qué manera. Tradicionalmente, los permisos para leer o escribir información se conceden si y sólo si se cumple alguna condición. Por ejemplo, el modelo de seguridad Bell-LaPadula establece que la información de alto nivel (es decir, secreta) no puede entrar en contenedores de bajo nivel (es decir, públicos).

Para aplicar esta política, cada contenedor de información (por ejemplo, un archivo) es clasificado (por ejemplo, secreto o público) en función del tipo de información que contiene y cada usuario tiene una autorización (por ejemplo, secreta

© WavebreakmediaMicro - Fotolia

o pública) en función del tipo de información que necesita conocer, dado su papel en la organización: un usuario de nivel secreto puede leer contenedores públicos o secretos, pero sólo puede escribir en contenedores secretos. Esto garantiza que la información secreta leída previamente nunca llegará a un contenedor público en el que un usuario de nivel público pueda leer.

En la práctica, la aplicación de una política de seguridad se basa generalmente en los mecanismos de control de acceso y control de flujo, implementados a nivel del sistema operativo.

4.2.1 Control de acceso

El control de acceso se refiere generalmente a la regulación de las solicitudes de acceso a los recursos gestionados por un sistema de información. Esta regulación puede producirse en varios lugares del sistema de información: a nivel de la red (en los firewall), a nivel de nodo (en los sistemas operativos) o a nivel de servicio (en las aplicaciones).

A nivel de red, los firewall permiten el acceso a los recursos de la red a los usuarios autenticados o al tráfico legítimo y deniegan el acceso a los usuarios no autenticados o al tráfico ilegítimo. Sin embargo, una mala configuración de un firewall puede provocar fallos de seguridad. En particular, las reglas de filtrado en conflicto pueden llevar a bloquear el tráfico legítimo y a aceptar paquetes no deseados.

En los sistemas operativos, el control de acceso aplica la política de seguridad mediante la autorización a sujetos (es decir, los usuarios o procesos autenticados) sobre objetos (es decir, los recursos como los datos, la red, las instalaciones de cálculo, etc.). Se han propuesto muchos modelos de control de acceso y algunos han sido ampliamente utilizados durante décadas, como los modelos DAC y RBAC. Estos modelos suelen representar las autorizaciones como tripletas (sujeto, objeto, permiso) y se emplean con frecuencia en sistemas operativos y bases de datos. En la última década se han propuesto muchas variantes de estos modelos, para capturar información adicional como el contexto del acceso (tiempo, ubicación, etc.), el propósito de las aplicaciones o las especificidades de las organizaciones; esto dio lugar al control de acceso basado en el contexto, en el propósito y en la organización.

En la actualidad, estas soluciones están bien establecidas y suelen estar normalizadas. En general, estos modelos consideran una arquitectura centralizada, es decir, los recursos y la información adicional necesaria para definir las políticas de control de acceso se encuentran en el lado del servidor y son gestionados por autoridades y administradores de confianza.

Más recientemente, la computación en la nube y las áreas emergentes de los sistemas centrados en el usuario y la Internet de las cosas (IoT), ponen una nueva luz en la investigación del control de acceso. Aunque el objetivo sigue siendo definir, evaluar y aplicar las autorizaciones, las especificidades intrínsecas de estos contextos inducen a un profundo replanteamiento de los modelos de control de acceso y

sus estrategias de aplicación. En la computación en la nube, los datos y los cálculos se subcontratan a entidades potencialmente no fiables, y la gestión del control de acceso debe adaptarse a nuevos supuestos de confianza. En los sistemas centrados en el usuario, la atención se centra en permitir la gestión del control de acceso sin recurrir a los expertos en TI. La IoT plantea el problema de adaptar el control de acceso cuando las Cosas, que disponen de escasos recursos, recopilan los datos.

Además, las Cosas (por ejemplo, los sensores integrados o los dispositivos GPS, los podómetros, los contadores inteligentes, los televisores conectados, los juguetes) registran los acontecimientos que se producen en su entorno y, por tanto, generan datos sensibles. Se están creando infinidad de nuevas aplicaciones y servicios que consultan estos datos. El diseño de modelos de control de acceso para la Internet de las Cosas es, por tanto, un problema crítico, aunque difícil, debido a dos objetivos conflictivos: (i) el modelo de control de acceso debe ser lo suficientemente genérico como para cubrir las necesidades de aplicaciones muy diversas y (ii) debe ser ligero teniendo en cuenta las limitaciones de hardware de las Cosas a las que se dirigen. Hasta ahora, los datos recogidos por las Cosas acaban en servidores centralizados donde se analizan y consultan.

En contextos centrados en el ser humano (por ejemplo, la nube personal), los individuos quieren gestionar ellos mismos sus datos personales, es decir, bajo su control y no delegar esta tarea en un administrador central. Sin embargo, el diseño de una política de control de acceso bien calibrada, y su aplicación, enfrenta a los individuos a la difícil elección entre delegar la administración de los datos a un tercero cualificado y renunciar a su soberanía o hacerse cargo de ella ellos mismos utilizando complejos modelos de intercambio y complicados protocolos de seguridad que probablemente no dominan. De ahí la necesidad de diseñar nuevos modelos de control de acceso, lo suficientemente sencillos como para ser gestionados por particulares.

En conclusión, aunque el control de acceso es un tema de investigación de larga duración, la evolución de las arquitecturas informáticas abre nuevas e importantes líneas de investigación. La tendencia actual sugiere que el control de acceso, habitualmente pensado en un contexto centralizado, debe considerarse en un contexto más distribuido y global, es decir, regulando los accesos a los datos a lo largo de todo el ciclo de vida de los mismos, desde las Cosas que los recogen hasta la Nube que los almacena. En términos de aplicación, esto también requiere complementar el control de acceso con técnicas de seguridad como el cifrado, el control de flujo de datos y la computación confiable.

4.2.2 Control de flujo de información

El control de flujo de información consiste en supervisar la forma en que una pieza de información fluye a través de un sistema o un programa. El objetivo es garantizar, ya sea de forma estática o dinámica, que un fragmento de información privada se manipule de acuerdo con su propiedad de seguridad. Una

propiedad típica es que una información privada no se filtre hacia un canal público.

SEGUIMIENTO DINÁMICO DEL FLUJO DE INFORMACIÓN

El seguimiento dinámico del flujo de información (DIFT por sus siglas en inglés) es una técnica popular y versátil para controlar el flujo de información a través de un sistema. Normalmente, los datos que entran en el sistema se etiquetan según sus niveles de seguridad y las etiquetas se propagan por el sistema. Esta técnica puede utilizarse para supervisar un sistema y garantizar que el flujo de información respeta la política de seguridad. Una propiedad básica es que los datos no confiables proporcionados por el usuario son saneados antes de ser procesados. El DIFT es también una poderosa herramienta para las pruebas de penetración y el análisis de vulnerabilidad: un error, por ejemplo, un desbordamiento del búfer, puede considerarse una vulnerabilidad de seguridad si los datos están influenciados por una entrada no fiable.

El análisis DIFT puede realizarse a varios niveles de granularidad. A nivel de sistema, el reto consiste en instrumentar el sistema operativo con etiquetas de seguridad y obtener una solución correcta, precisa y que requiera un mantenimiento limitado. Para mejorar la escalabilidad de DIFT, se han propuesto varias soluciones de hardware. Por ejemplo, el proyecto CrashSafe propone un novedoso diseño de procesador. También hay propuestas menos intrusivas en las que un coprocesador independiente se dedica a gestionar las etiquetas de seguridad.

La DIFT suele limitarse a los flujos de información directa que pueden aplicarse mediante la supervisión de una ejecución a la vez. Notablemente, DIFT también puede hacer cumplir hiper-propiedades, como la no interferencia, que son propiedades de conjuntos de trazas. Como esta supervisión detiene la ejecución en caso de violación, convierten una fuga de información en una fuga de terminación que puede ser aceptable para algunas aplicaciones.

SEGUIMIENTO ESTÁTICO DEL FLUJO DE INFORMACIÓN

Existe una amplia literatura sobre sistemas de tipos para el flujo de información. Los sistemas de tipos sofisticados manejan características ricas del lenguaje como las excepciones y el envío de métodos. El sistema JIF para Java es probablemente la implementación más impresionante de control estático de flujo de información. La propiedad tradicional que se impone en los sistemas de tipos es la no interferencia, que básicamente establece que los datos públicos no dependen de los datos privados. Para acomodar los programas que filtran una cantidad controlada de información, por ejemplo, las primitivas criptográficas, los sistemas de tipos prácticos necesitan incluir una noción de desclasificación. La desclasificación puede adoptar varias formas, pero requiere algún tipo de especificación proporcionada por el usuario en la que el sistema confíe. En general, es difícil predecir si el efecto a largo plazo de una desclasificación se corresponde

con la intención del usuario.

FLUJO DE INFORMACIÓN CUANTITATIVO

Una alternativa a la desclasificación es el flujo de información cuantitativo (QIF, o Quantitative Information Flow en su denominación inglesa), cuyo objetivo es cuantificar de forma estática o dinámica la cantidad de información que se filtra por un sistema, a través de alguna observación del atacante.

Para evitar el QIF, los métodos típicos de seguridad, como el cifrado y el control de acceso, no son aplicables: la única manera es ofuscar el vínculo entre el secreto y el observable. Lo ideal sería que los sistemas fueran completamente seguros, pero en la práctica este objetivo es a menudo imposible de alcanzar. Por lo tanto, es importante disponer de una teoría cuantitativa de las fugas, para medir la vulnerabilidad de un secreto, valorar si un sistema es mejor que otro o evaluar la efectividad de un método para evitar las fugas. Los aspectos cuantitativos se derivan del hecho de que el conocimiento del adversario es típicamente de naturaleza probabilística y que los mejores métodos para prevenir la filtración son a menudo métodos aleatorios. Los enfoques más exitosos de los fundamentos del flujo de información cuantitativo se basan en la teoría de la información y en la noción de entropía (en varias versiones: Shannon, Rényi o adivinación). La entropía mide la vulnerabilidad del secreto y la elección entre varias nociones refleja los diferentes modelos operativos de adversarios que a uno le interesan.

Un primer inconveniente de QIF es que las distintas definiciones de filtración ofrecen garantías de seguridad diferentes, a veces incomparables. El enfoque de QIF abarca una gran variedad de ataques (conjeturas aproximadas, conjeturas múltiples o propiedades del secreto) y subsume la mayoría de los enfoques teóricos de la información considerados en la literatura y sus correspondientes nociones de entropía.

Un segundo problema de QIF es que las garantías probabilísticas medias puede que no proporcionen una garantía de seguridad adecuada en caso de un atacante activo, es decir, cuando el atacante controla algunos de los datos de entrada y trata de utilizarlos para activar vulnerabilidades de seguridad.

COMPILADORES PARA ASEGURAR LA APLICACIÓN DEL FLUJO DE INFORMACIÓN

Otra tendencia de investigación tiene como objetivo preservar las propiedades del flujo de información de alto nivel a través de la cadena de compiladores. El flujo de información es importante para los lenguajes de varios niveles en los que el compilador tiene la responsabilidad de decidir qué nivel almacena la información sensible y cómo garantizar la seguridad de las comunicaciones. El flujo de información también es un asunto importante para el código crítico, por ejemplo, las primitivas criptográficas, donde cualquier fuga debida a la implementación puede romper la garantía de seguridad matemática. En este contexto, el

compilador tiene como objetivo proteger contra los ataques de canal lateral, como los ataques de sincronización de información o de análisis del consumo de energía.

[Equipos Inria] Control de acceso y control de flujo

↗ El equipo **CELTIQUE** desarrolla técnicas y análisis de compiladores certificados para la protección contra ataques de sincronización de información. Junto con el equipo **INDES**, **CELTIQUE** desarrolla una teoría para monitores híbridos que aumenta un monitor dinámico de flujo de información con un análisis estático (dinámico) de ramas no ejecutadas calculando una forma simbólica y cuantitativa del flujo de información. Junto con el equipo **MARELLE**, **CELTIQUE** trabaja brindando soporte para compiladores para la programación de tiempo constante, una estricta disciplina de programación adoptada por los criptógrafos para limitar las fugas de ataques de sincronización.

↗ El equipo **CIDRE** tiene una gran experiencia en el monitoreo de Flujos de Información Dinámicos tanto a nivel de sistema como de hardware: el equipo desarrolla la herramienta blare, un sistema de detección de intrusiones (IDS, por sus siglas en inglés) que permite, en Linux y Android, evaluar dinámicamente la legalidad de los flujos de información. Se ha diseñado un dispositivo de hardware para mejorar la precisión de esta evaluación.

↗ El equipo **COMETE** propuso la teoría para fundamentar la seguridad en nociones de flujo de información cuantitativo y desarrolló el marco de trabajo g-leakage. También desarrollaron la biblioteca Libqif [a], un conjunto de herramientas en C++ que implementa diversas técnicas relacionadas con g-leakage, flujos de información cuantitativa y privacidad diferencial. Este equipo también está investigando un enfoque basado en la Teoría de Juegos para limitar la fuga de información en presencia de un atacante activo.

↗ El equipo **FUN** investiga una forma de evitar la transmisión de todos los datos recogidos por las Cosas a un servidor. El equipo utiliza almacenamiento descentralizado para evitar reconstruir todos los datos a la vez. Además, identifica los nodos de retransmisión maliciosos para evitarlos durante el proceso de recolección de datos.

↗ El equipo **INDES** trabaja en la definición, comparación y evaluación de las limitaciones de diferentes políticas de flujo de información –incluyendo la desclasificación y la no interferencia computacional para la criptografía– a nivel de lenguaje. Los dominios de las aplicaciones incluyen principalmente JavaScript, lenguajes reactivos y multinivel para la IoT. El equipo también desarrolla mecanismos de sonido estáticos, dinámicos e híbridos para forzar la seguridad del flujo de información y trabaja en preservación por compilación de garantías del flujo de información.

↗ El equipo **RESIST**, en colaboración con el equipo **PESTO**, trabaja en métodos de gestión de los archivos de configuración de firewall que revelan automáticamente anomalías y ayudan al administrador a encontrar una posición adecuada para una regla de filtrado recién añadida.

↗ El equipo **PETRUS** diseña mecanismos de control de acceso para la nube personal que no requieren que el usuario humano entienda los mecanismos de control de acceso

subyacentes para aplicar una determinada política de seguridad. En el contexto de la IoT, el equipo diseña modelos de control que se apoyan en estructuras de gestión de datos y algoritmos integrados para las Cosas, de modo que las decisiones de difusión de datos puedan evaluarse más cerca de los datos, dentro de las Cosas que los recopilaron.
↗ El equipo **PROSECCO** PROSECCO tiene como objetivo diseñar compiladores formalmente seguros para la arquitectura con soporte para el control dinámico del flujo de la información. El equipo también estudia cómo pueden aprovecharse las soluciones de hardware para el control dinámico del flujo de la información con el fin de garantizar que el código en ejecución satisfaga las políticas de seguridad.
↗ El equipo **VALDA** estudia tanto los aspectos fundacionales como los sistémicos de la gestión de datos complejos, especialmente los centrados en el ser humano. Propusieron un modelo de control de acceso colaborativo en el contexto del lenguaje Webdamlog (datalog distribuido). Este modelo permite a los individuos especificar de forma manifiesta potentes políticas que rijan el acceso a sus datos, la difusión de los mismos y la delegación de su procesamiento.

a. https://github.com/chatziko/libqif

4.3 Computación confiable

[Resumen]
La computación confiable se basa en las garantías que ofrece el hardware seguro. Este hardware puede asegurar la integridad de la plataforma y un arranque sin riesgos. También es posible proporcionar confianza a nivel de aplicación en lugar de a nivel de toda la plataforma, ejecutando el código entornos de actuación aislados y confiables, llamados enclaves. Además, estos enclaves ofrecen la posibilidad de certificar que los resultados han sido producidos por un código determinado, lo que abre la posibilidad de externalizar el cálculo de tareas .

La computación confiable consiste en construir hardware seguro para obtener garantías globales sobre los cálculos realizados en una plataforma. El primer uso de este enfoque tenía como objetivo garantizar su integridad mediante el uso de módulos de plataforma confiable (TPM, o Trusted Platform Module por su denominación en inglés) y un arranque seguro. Un TPM incluye una clave criptográfica única y puede calcular funciones hash. Esto le permite autenticar los dispositivos de hardware y verificar que el software no ha sido modificado, y así certificar la integridad de toda la secuencia de arranque. De este modo, el usuario puede garantizar a terceros que su máquina está ejecutando un sistema operativo (SO) y unas aplicaciones concretas. Este enfoque tiene como objetivo proporcionar una integridad total de la plataforma bajo

el supuesto de que el sistema operativo y las aplicaciones son de confianza. Siguiendo este enfoque, se han desarrollado monitores de máquinas virtuales (VMM, Virtual Machine Monitors) basados en TPM, que permiten el aislamiento de múltiples SO ajenos y, por tanto, el aislamiento de los ataques a estos SO. Esto condujo al desarrollo de micronúcleos (por ejemplo, SEL4) y uninúcleos (por ejemplo, MirageOS), con el objetivo de minimizar la confianza requerida en el SO y sus aplicaciones.

Más recientemente, con el aumento de las zonas seguras en los procesadores principales (por ejemplo, el Software Guard Extension (SGX) y Trustzone de ARM), se puede garantizar que el código y los datos de la memoria están protegidos con respecto a la confidencialidad y la integridad. Es así posible garantizar la confianza a nivel de una aplicación concreta en lugar de a nivel de toda la plataforma. De hecho, estos entornos de ejecución aislados ofrecen la posibilidad de ejecutar código en enclaves, cuya memoria y flujo de control están protegidos del entorno (incluido el sistema operativo). Además, estos enclaves proporcionan funciones de certificación, es decir, medios para que una parte externa compruebe que los mensajes presentados se produjeron realmente en un enclave que ejecuta una pieza de código específica. Este enfoque mucho más versátil conduce a diversas líneas de investigación interesantes para asegurar los cálculos tanto en la nube como en los dispositivos domésticos. Entre las aplicaciones típicas se encuentra la concesión de licencias seguras, que aprovecha las garantías proporcionadas por los enclaves para asegurar que el software con licencia no se utiliza de forma ilegal.

Un reto importante relacionado con esta tecnología disruptiva es el estudio de su aplicabilidad en el caso de la externalización segura de cálculos de datos (posiblemente dispersos). Dado que los Entornos de Ejecución Confiables (Trusted Execution Environments, o TEE) proporcionan garantías de seguridad en el hardware como la confidencialidad, la integridad y la certificación, el código que se ejecuta dentro de un enclave TEE (local o remoto) puede considerarse que se acerca a un comportamiento totalmente fiable. Esto abre el camino a tareas de computación genéricas, eficientes, escalables y seguras orientadas a los datos. VC3 es un típico intento preliminar de Microsoft para conseguir una analítica de datos confiables basada en enclaves Intel SGX en la nube. Sin embargo, hay que superar dos problemas principales en la investigación. En primer lugar, las propiedades de seguridad del hardware proporcionadas por los enclaves TEE no pueden considerarse incondicionalmente inviolables y deben llevar a investigar modelos de amenaza ligeramente diferentes. En segundo lugar, las tareas más comunes orientadas a los datos (por ejemplo, la búsqueda privada en la web, el procesamiento seguro de flujos de datos en la IoT, el aprendizaje automático que preserva la privacidad) deben transformarse eficientemente en contrapartidas seguras basadas en TEE. El reto aquí es optimizar y asegurar la ejecución de las funciones primitivas orientadas a los datos de acuerdo con las restricciones de los TEE.

Estas líneas de investigación están surgiendo hoy en día. La tendencia actual sugiere que la disponibilidad y la diversidad de las tecnologías TEE aumentarán en un futuro

próximo. Se avecinan nuevas soluciones, con plataformas multinúcleo en las que los núcleos orientados a la seguridad y el aislamiento (a la manera de SGX) cohabitarán con otros núcleos polivalentes, permitiendo así la separación de tareas dentro de la CPU. Esto sugiere que la computación confiable será cada vez más importante en un futuro próximo, e Inria contribuirá a los esfuerzos de investigación en esta dirección.

[Equipos Inria] Computación confiable

↗ El equipo **PETRUS** trabaja en nuevas propiedades para cuestionarse sobre el cómputo seguro del hardware en el contexto del cómputo de bases de datos dispersas, por ejemplo, limitando la cantidad de datos accesibles a cada agente/enclave. El equipo también estudia el procesamiento de bases de datos utilizando elementos seguros o TEEs más avanzados, como Intel SGX.

4.4 Detección de intrusos y correlación de alertas

[Resumen]

Casi cualquier sistema contiene, generalmente de forma involuntaria, defectos. Un atacante puede aprovecharlos para saltarse los mecanismos de seguridad preventiva existentes, por ejemplo, el control de acceso. Por lo tanto, la supervisión del sistema es de crucial importancia para identificar cualquier violación de la política de seguridad. Para detectar intrusiones, el enfoque principal y más extendido consiste en definir los indicadores maliciosos y buscar su aparición en diversas fuentes de información (registros de la red, del sistema operativo y de las aplicaciones, etc.). Una alternativa consiste en definir las actividades normales del sistema monitorizado y medir las posibles desviaciones de esta normalidad. En ambos casos, el reto consiste en detectar todas las intrusiones, pero sólo las intrusiones. Sin embargo, en la práctica, la detección dista mucho de ser tan perfecta, dando lugar a numerosos falsos positivos (falsas alertas) o falsos negativos (ataques no detectados). Por ello, varios equipos de Inria exploran distintas formas de generar alertas. La detección de intrusos da lugar a un gran número de alertas, muchas de las cuales son falsos positivos. Por ello, es necesario un paso adicional: la correlación de alertas. Este paso consiste en aplicar en el proceso de alerta una serie de transformaciones para mejorar progresivamente su contenido (por ejemplo, añadiendo información sobre el éxito de los ataques correspondientes, sobre el origen de los ataques, sobre la vulnerabilidad que ha sido explotada, sobre las alertas relacionadas, etc.) y aumentar así el "conocimiento de la situación" del administrador.

Numerosos defectos (es decir, vulnerabilidades) se producen en cualquier sistema durante su diseño, implementación o configuración. Por lo general, esta presencia no es intencionada, pero la introducción maliciosa de vulnerabilidades

sigue siendo una posibilidad. Estos puntos débiles pueden ser explotados para eludir los mecanismos de seguridad preventiva utilizados para hacer cumplir la política de seguridad. Además de la seguridad preventiva, es obligatoria una segunda línea de defensa, la detección de intrusos. Consiste en vigilar los sistemas para detectar cualquier violación de la política de seguridad aplicada. Por "intrusión" se entiende la "violación de la política de seguridad" que rompe la confidencialidad, la integridad o la disponibilidad. El uso de un malware (virus, gusano, bomba lógica, etc.) es, por supuesto, una buena manera de llevar a cabo dicha violación.

La detección de intrusiones es un servicio de seguridad reactivo que consiste en recoger información sobre el funcionamiento del sistema vigilado y analizar estas actividades para producir alertas si se consideran maliciosas. Como el analizador es muy a menudo propenso a generar falsas alertas, una segunda etapa de análisis evalúa el flujo de alertas e intenta eliminar las falsas. Además, las alertas se correlacionan para identificar escenarios de intrusión secuencial en varias fases. Por último, una vez que se puede confiar en las alertas existentes, se puede considerar la reacción a las intrusiones detectadas.

4.4.1 Paradigmas de detección de intrusos

Un sistema de detección de intrusiones (IDS por sus siglas en inglés) analiza los datos procedentes tanto del tráfico de la red (Network-Based IDS, NIDS), del sistema operativo o de las aplicaciones (Host-Based IDS, HIDS).

El análisis de estas dos categorías de datos sigue dos paradigmas de detección diferentes: el reconocimiento de intrusiones basado en el uso indebido y el basado en las anomalías. La detección de anomalías consiste en definir las actividades normales de la entidad vigilada y en identificar cualquier desviación de esta normalidad; mientras que la detección de usos indebidos consiste en modelizar las actividades maliciosas y detectar las frecuencias de estas actividades.

El enfoque más clásico de detección de usos indebidos, popularizado a finales de los años 90 con el NIDS snort[82], consiste en buscar rastros de ataques conocidos en los paquetes de red. Implica la actualización constante de una base de datos de registros de ataques; los 'ataques de día cero'[83] no suelen detectarse. Además, hay que llegar a un balance entre la selectividad de los registros y el riesgo de falsos positivos y falsos negativos. Los registros muy selectivos aumentan el riesgo de omitir variantes del ataque, mientras que los más genéricos pueden dar lugar a un alto índice de falsas alarmas. En la práctica, los registros suelen ser muy sencillos y, por lo tanto, muy genéricos, para permitir el análisis en tiempo real de cada evento, lo que suele dar lugar a una mayor tasa de falsos positivos.

La detección de anomalías es menos frecuente que la detección de usos

82. https://www.snort.org/
83. Una vulnerabilidad de día cero es aquella que es desconocida para aquellos que estarían interesados en mitigarla.

indebidos. Aquí, definir la normalidad mediante un modelo de comportamiento es, por supuesto, el punto crítico. Si el modelo es demasiado preciso, el detector genera un elevado número de falsas alarmas; si es demasiado laxo, pasa por alto los ataques. Encontrar un buen equilibrio es difícil, especialmente cuando el modelo se construye utilizando la estadística o el aprendizaje automático, debido a la posibilidad de sub-entrenamiento o sobre-entrenamiento. Además, un buen equilibrio para una actividad determinada puede resultar inaceptable más adelante cuando la actividad evoluciona. En cualquier caso, tratar con un comportamiento legítimo pero no planificado es problemático para la detección de anomalías.

La información recogida en los paquetes de red puede ser a veces demasiado pobre desde el punto de vista semántico para permitir un buen proceso de detección. Por ello, la detección de intrusiones y anomalías debe abordarse también tanto a nivel de aplicación como de sistema operativo. Los mecanismos de detección de anomalías más comunes a nivel de aplicación consisten en detectar una desviación del flujo de control de un programa. Un método popular para detectar dicha anomalía es el uso de secuencias de aplicación de llamadas al sistema. Sin embargo, estos métodos pueden ser eludidos por ataques de imitación[84][WS02] o por ataques contra la integridad de los parámetros de las llamadas al sistema.

[Desafío de investigación 6] Detección de intrusos en redes encriptadas

Hoy en día, la detección de intrusiones se realiza esencialmente a nivel de la red. Si, como se espera en un futuro próximo, el tráfico se encriptara más sistemáticamente, lo que por supuesto sería una buena práctica para la seguridad y la privacidad, el análisis de los paquetes de red se volvería de facto inoperante, por no hablar del análisis del encabezado de los datos. Por lo tanto, se hace importante estudiar y diseñar nuevos mecanismos de supervisión de los sistemas de información y de producción de alertas, a nivel de aplicación, de software de intercambio (middleware), de sistema operativo, e incluso de firmware o de hardware.

4.4.2 Correlación de alertas

Se distingue entre los enfoques de correlación de alertas implícitas y explícitas. La correlación implícita de alertas utiliza paradigmas de minería de datos para fusionar o agregar alertas, basándose simplemente en la similitud entre las características de las alertas (por ejemplo, las direcciones IP de la víctima y del atacante), o utilizando técnicas más avanzadas para extraer información relevante de los grupos de alertas (minando reglas entre sus atributos). La correlación explícita de alertas se basa en que los expertos en seguridad especifiquen las restricciones lógicas y temporales entre las alertas que corresponden a escenarios de ataque

84 [WS02] David Wagner y Paolo Soto. Mimicry attacks on host-based intrusion detection systems. En Proceedings of the 9th ACM conference on Computer and communications security (CCS), páginas 255-264, New York, NY, USA, 2002. ACM.

complejos, que generalmente requieren varios pasos para lograr su objetivo final. Cuando se detecta un escenario de intrusión completo o parcial, se genera una alerta de nivel superior.

La información adicional sobre las características de los ataques y sobre el contexto en el que se producen también es útil para el proceso de correlación. Este conocimiento permite tener en cuenta el contexto al procesar las alertas, para identificar falsos positivos o para evaluar si un determinado ataque tiene alguna posibilidad de éxito dado el contexto en el que se produce.

Por desgracia, a menudo sigue siendo difícil escribir reglas de correlación que aprovechen correctamente toda la información disponible y que traduzcan de forma adecuada los conocimientos del administrador del sistema en relación con los posibles ataques contra el mismo. Por ello, la automatización de la producción de reglas es un tema de investigación actual.

[Equipos Inria] Detección de intrusos y correlación de alertas

↗ El equipo **CIDRE**ha estudiado ampliamente la detección de intrusiones (a nivel de aplicación y de red) y la correlación de alertas. A nivel de aplicación, el equipo ha propuesto un enfoque para detectar la corrupción de elementos de datos que tienen una influencia en las llamadas al sistema. Este enfoque consiste en construir automáticamente un modelo de comportamiento orientado a los datos de una aplicación mediante el análisis estático de su código fuente que se utiliza para construir restricciones sobre los datos manipulados por el programa. A continuación, la aplicación se instrumenta con aserciones ejecutables para comprobar estas restricciones en tiempo de ejecución. A nivel del sistema operativo, el equipo propuso un enfoque de detección de anomalías en el que el modelo de comportamiento no se aprende, sino que se da en forma de una política de flujo de información. La idea básica es definir dónde puede ser almacenada cada información (por ejemplo, la información contenida en cada archivo en la inicialización del sistema), potencialmente mezclada con otra. En cuanto a la correlación, el equipo introdujo un lenguaje de descripción de ataques que permite definir las alertas que debe producir el ataque y las reglas lógico-temporales entre estas ellas. Estas reglas se utilizan para configurar un motor de correlación realizado por el equipo. **CIDRE** también propuso con sus colaboradores un modelo de datos para representar un sistema bajo vigilancia: este modelo puede utilizarse durante el proceso de correlación para aportar información contextual, por ejemplo para la identificación de falsas alertas. Como estas reglas de correlación son a veces difíciles de definir, el equipo también propuso, en colaboración con la DGA, un proceso exhaustivo para generar dichas reglas de la forma más automática posible, partiendo de árboles de ataque (véase §3.4) que normalmente ya han sido detectados por el administrador para evaluar las amenazas contra su sistema durante la fase de análisis de riesgos.

↗ El equipo **LACODAM** aborda el análisis de grandes datos de red para identificar posibles amenazas persistentes avanzadas (APT o advanced persistent threats por su acepción inglesa) descubriendo patrones sintomáticos en los metadatos de los paquetes IP. Encontrar estos complejos patrones en un gran volumen de datos de streaming implica revisar los algoritmos de minería de flujos existentes para mejorar su rendimiento de forma drástica, garantizando al mismo tiempo una tasa de falsos positivos manejable.

↗ El equipo **MYRIADS** investiga la detección de usos indebidos en contextos de la nube. Esto es especialmente complicado, ya que el sistema de información puede reconfigurarse de forma dinámica y automática. Los mecanismos de supervisión de la seguridad deben estar bajo el control del proveedor de la nube y deben seguir la dinámica del entorno. En este contexto, el equipo propuso un sistema de detección de usos indebidos autoadaptable para nubes IaaS, que supervisa los cambios en la infraestructura virtual de un entorno de nube y reconfigura las sondas de seguridad en consecuencia. Además, el equipo propuso un método que permite a un cliente de la nube verificar que un sistema de detección de intrusiones en la red situado en la infraestructura del operador está correctamente configurado, de acuerdo con los objetivos de nivel de servicio incluidos en el Acuerdo de Nivel de Servicio (SLA).

↗ Centrándose en los datos de la red, el equipo **RESIST** trabaja en la creación de soluciones para caracterizar y detec tar comportamientos no deseados en la red. El equipo propuso un método de agrupación y visualización que permite analizar un gran número de paquetes IP para que los analistas de seguridad puedan observar fácilmente los patrones de actividad maliciosa. El equipo también propuso una técnica para investigar el tráfico https (por tanto, cifrado). El equipo definió características específicas para el tráfico https que se utilizan como entrada para los algoritmos de aprendizaje automático que procesan sesiones tls completas. Esto permite la identificación temprana de servicios web encriptados en la sesión tls con un alto grado de precisión, lo que luego posibilita la detección de anomalías en el uso de los servicios identificados. Otra contribución interesante del equipo está relacionada con la cuantificación del número de nodos de monitorización necesarios para garantizar una tasa de falsos positivos aceptable para diferentes topologías de red. El equipo ha demostrado que la tasa de falsos positivos puede reducirse mediante la colocación estratégica de los nodos de vigilancia.

4.5 Análisis y detección de malware

[Resumen]

El malware (virus, gusanos, ransomware, spyware, adware, troyanos, keyloggers, rootkits, etc.) es, por supuesto, una gran amenaza para nuestros sistemas de información (SO, aplicaciones y datos), especialmente en el lado del cliente (PC o smartphones).

El objetivo del análisis de malware es obtener una comprensión completa de un presunto código malicioso: identificar los objetivos (por ejemplo, un usuario final en particular, una máquina que ejecuta un sistema operativo concreto), las acciones de ataque (por ejemplo, la fuga de información, o el cifrado y el rescate), las técnicas para eludir los mecanismos de seguridad, y sus propios mecanismos de protección para evitar la detección. Para tener éxito, el análisis debe vencer primero las protecciones anti-análisis establecidas por el creador del malware (ofuscación).

La detección de malware suele basarse en el análisis de cualquier información recibida por un dispositivo (una máquina, un teléfono, un firewall) o incluso en el escaneo completo de los archivos contenidos en una máquina. El motor de detección compara los datos recuperados con una base de datos de características conocidas y sintomáticas de sufrir malware (es decir, firmas de malware). El reto consiste en mantener una base de datos de firmas actualizada, ya que los autores de malware generan constantemente nuevas versiones basadas en el mismo código malicioso para escapar al escrutinio. Por ello, los proyectos de investigación proponen técnicas de detección basadas en el comportamiento concreto del malware, que se mantiene constante en todas las versiones.

4.5.1 Análisis de malware

El análisis de malware pretende diseccionar cualquier pieza de código identificada como sospechosa y potencialmente maliciosa, con el objetivo de comprender plenamente el código malicioso para mejorar los mecanismos de seguridad existentes o diseñar nuevas contramedidas. En concreto, el análisis de malware tiene como objetivo identificar sus objetivos (un usuario final concreto, una empresa, cualquier máquina con un sistema operativo específico, etc.), las acciones que pretende realizar para atacar a los objetivos (filtrado de información sensible, cifrado y rescate, etc.), la forma en que consigue eludir los mecanismos de seguridad que protegen a los objetivos y la forma en que se protege a sí mismo contra los motores de detección.

Dado el impacto potencial de un malware, es crucial que el análisis se realice lo más rápidamente posible. Con respecto a esta perspectiva, las contribuciones científicas sobre el análisis de malware tienen dos vertientes. En primer lugar, algunos enfoques se centran en la clasificación automática. Su objetivo es distinguir el código benigno del malicioso y, a continuación, clasificar el código malicioso en una de las familias conocidas. Se han realizado importantes esfuerzos en este ámbito que han

permitido disminuir la carga de trabajo humana al reducir el número de muestras que hay que analizar manualmente. Las muestras restantes suelen proceder de malware desconocido o están demasiado protegidas para ser procesadas automáticamente. Así, otros enfoques pretenden ayudar a los expertos a realizar ingeniería inversa y comprender el código malicioso.

Se puede realizar un análisis estático, sin ejecutar el malware. Esto ofrece varios beneficios. En primer lugar, el análisis estático es seguro para la arquitectura del cliente, ya que el código malicioso no se ejecuta. En segundo lugar, el análisis estático aporta información sobre todas las posibles ejecuciones del código y, por tanto, sobre todos los posibles comportamientos del malware.

Por desgracia, como los autores de malware son conscientes de que su código probablemente se enfrentará a un análisis estático, utilizan varias técnicas para dificultar dicho análisis y su ingeniería inversa. Utilizan técnicas de ofuscación como el empaquetado, la atenuación del flujo de control o los predicados opacos. Todas estas técnicas hacen que el gráfico de flujo de control calculado a partir del código malicioso sea irrelevante para la mayoría de las herramientas de análisis estático.

El análisis también puede realizarse de forma dinámica, lo que significa que el código malicioso se ejecuta en la medida de lo posible. El objetivo principal de este tipo de análisis es observar los efectos reales y concretos del ataque en su objetivo. Una dificultad en este caso es provocar la ejecución de la parte maliciosa del código.

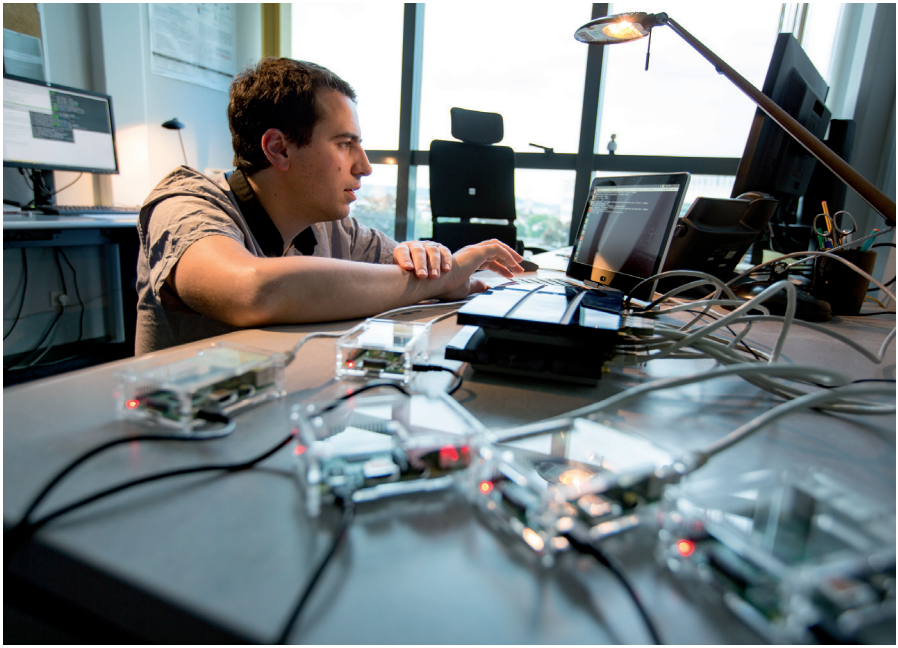
4.5.2 Detección de malware

Aquí distinguimos dos tipos de malware. Los programas maliciosos normales de tipo comercial pretenden causar daños a una población importante de usuarios finales lo más rápidamente posible y pueden encontrarse en plataformas populares como Windows o Android. En cambio, el malware específico para un objetivo pretende causar daños en plataformas concretas, como el gusano Stuxnet, que se dirigió a los controladores lógicos programables de las centrales nucleares iraníes. El malware específico para un objetivo se construye con gran cuidado y, por tanto, es especialmente difícil de detectar[85]. Además, al ser más sigilosos, suelen ser identificados tardíamente, cuando los síntomas externos revelan su existencia: Stuxnet fue revelado en 2010, pero se cree que fue desplegado en 2005 o incluso antes.

Los programas maliciosos habituales de tipo comercial se desarrollaron inicialmente y en su mayoría para atacar los sistemas operativos y las aplicaciones de Windows, ya que era el sistema operativo más utilizado. Sin embargo, durante los últimos diez años, Android también se ha convertido en un objetivo popular para los autores de malware. Un programa malicioso puede infectar a su objetivo desde diferentes

85. El antiguo equipo de Carte dirigió un análisis profundo de Duqu (https://en.wikipedia.org/wiki/Duqu), uno de los malware más sofisticados de 2011, y desarrolló un detector de malware basado en el análisis morfológico, una técnica basada en la comparación de gráficos de flujos de control. La actividad de análisis de malware se desarrolla en el equipo Carbone de LORIA.

puntos de entrada, como aplicaciones descargadas por el usuario, sitios web visitados o archivos adjuntos de correo electrónico, y puede propagarse a través de diferentes redes. La detección del malware suele producirse durante el escaneo del tráfico de la red (archivos adjuntos al correo, por ejemplo) o de las máquinas. En ambos casos, el motor de detección compara los datos a analizar con una base de datos de firmas de malware (características conocidas que son sintomáticas de malware). El principal reto consiste en construir y mantener actualizada una base de datos registrada.



Detección de comportamientos no deseados en la red – © Inria / Foto C. Morel

Se trata de un juego del gato y el ratón: por un lado, los autores de programas malignos intentan evitar la detección el mayor tiempo posible mientras minimizan el esfuerzo de código de producción y, por otro lado, los defensores tienen que producir registros de malware lo más precisos y completos posibles mediante un análisis profundo del malware conocido, como se explica en la sección anterior.

Los autores de malware suelen basarse en paquetes de malware y técnicas de blindaje adicionales para generar nuevos archivos ejecutables basados en el mismo código malicioso. De este modo, se puede generar un gran número de variantes a partir de una única muestra original. Por ejemplo, cada 15 segundos se generan nuevas variantes de Cerber. La detección de un nuevo malware o de una nueva variante de un malware conocido sigue siendo un problema no resuelto. Un avance reciente es

el uso de enfoques de comportamiento, en lugar de firmas estáticas, mediante el seguimiento de cualquier desviación de un modelo del comportamiento normal, por ejemplo, mediante el modelado del flujo de información a nivel del sistema operativo.

A partir de 2012, las estafas basadas en ransomware han crecido a nivel internacional. Un ransomware es un tipo específico de malware que restringe el acceso a un sistema informático o a sus datos alojados (clásicamente cifrando los datos). Un ransomware requiere que el usuario pague un rescate al atacante para eliminar la restricción. La aparición de los ransomware está probablemente relacionada con el sistema Bitcoin, que permitía los pagos anónimos y hacía rentables los ataques masivos.

[Equipos Inria] Análisis y detección de malware

↗ El equipo **CIDRE** trabaja en el análisis de malware, la detección de malware y la desofuscación de malware. En cuanto al análisis de malware, su objetivo es el entorno Android. La idea es utilizar el rastreo de flujos de información para analizar el comportamiento del malware y, potencialmente, generar un registro del malware basado en sus actividades reales (es decir, los flujos de información generados). Como todo análisis dinámico, este trabajo sólo es relevante si la parte maliciosa del código se ejecuta realmente. Así, el equipo se centra en la activación automática del código malicioso. La herramienta de detección de malware supervisa el uso de los datos del sistema de archivo para comprobar las desviaciones de los datos con respecto a su uso normal. Además, una acción postmortem mediante un algoritmo de aprendizaje automático no clasificado proporciona pistas para identificar claramente el malware detectado. El trabajo de desofuscación de malware tiene como objetivo eludir las protecciones del propio malware contra el análisis estático.
El LHS Rennes recopila ransomware (plataforma Malware'O'Matic). Como este tipo de malware tiene una vida corta, se requiere renovar periódicamente la base de datos y verificar qué ransomware siguen 'vivos'.

4.6 Reacción a los ataques detectados

[Resumen]

Idealmente, la detección de intrusiones y malware, así como la correlación de alertas, deberían conducir a la detección de todos los ataques sin falsas alertas. Por lo tanto, el siguiente paso obvio es responder (posiblemente de forma automática) a los ataques detectados mediante acciones apropiadas: modificación de la política de seguridad, nuevas configuraciones de los mecanismos de seguridad existentes, implementación de nuevos mecanismos de seguridad, despliegue de parches, etc. Por supuesto, es importante evitar que las contramedidas tengan consecuencias similares o incluso peores que las del propio ataque.

Idealmente, la sincronización de alertas debería conducir a la detección de todos los ataques sin falsas alarmas. Por lo tanto, el siguiente paso obvio debería ser como es lógico responder (potencialmente de forma automática) a los ataques detectados. La respuesta automática también se hace necesaria por la rapidez con la que se puede propagar un ataque, y el considerable daño que podría hacer antes de la respuesta manual. Si se considera que la política de seguridad ha sido violada aunque se hayan utilizado mecanismos preventivos para hacerla cumplir, se pueden considerar dos niveles de reacción: (1) el ataque puede haber tenido éxito porque la política era incorrecta, en cuyo caso la política debe ser modificada, y se deben establecer nuevas configuraciones de los mecanismos de seguridad existentes o incluso nuevos mecanismos de seguridad; (2) el ataque también puede haber tenido éxito porque la aplicación de la política era incorrecta, en cuyo caso deben identificarse y corregirse los errores de configuración del mecanismo de seguridad. El uso de métodos formales puede ayudar a garantizar que las propiedades de seguridad solicitadas por la política están efectivamente aseguradas a nivel de política y a nivel de aplicación. Por supuesto, es importante que las contramedidas (modificación de la política, nuevas configuraciones, nuevos mecanismos de seguridad, parches, etc.) no sean similares o peores que las del propio ataque. Por ejemplo, al intentar detener un ataque DDoS, los paquetes legítimos también pueden ser rechazados: el servicio quedará entonces no disponible para los usuarios legítimos, que era el objetivo último del ataque.

[Equipos Inria] Reacción a los ataques detectados

↗ El equipo **CTRL-A** se centra en el tema relativamente poco estudiado de los modelos y técnicas de control para la reacción automatizada a los ataques. El equipo utiliza la información de detección para identificar las acciones de defensa y reparación apropiadas, de modo que el sistema pueda seguir siendo operativo, completamente o en modo reducido. En términos de Computación Autónoma, esta capacidad se denomina autoprotección.

↗ Como reacción al ataque DDoS, los paquetes legítimos pueden ser rechazados. El servicio puede entonces no estar disponible para los usuarios legítimos. Por ello, el equipo **RESIST** propone separar el tráfico del DDoS en el tiempo y el espacio: al introducir retrasos voluntarios y rutas más largas, aunque el rendimiento puede degradarse, el servicio no se interrumpe.

5

Privacidad y protección de datos personales

La privacidad se suele definir como la capacidad de los individuos de controlar sus datos personales y decidir qué revelar, a quién y bajo qué condiciones. Sin embargo, no existe una única definición, ya que la noción de privacidad está íntimamente ligada a nuestras raíces culturales. Por ejemplo, la noción de datos personales, piedra angular de la privacidad, ha recibido varias acepciones, no siempre compatibles, dependiendo de los países. En Francia, los datos personales se incluyen en la "Ley de Protección de Datos" de 1978, como la información que puede vincularse directa o indirectamente a una persona, por parte del responsable del tratamiento de los datos, o de cualquier tercero, utilizando cualquier tipo de medio. Por lo tanto, tiene un alcance muy amplio, independientemente de la naturaleza de los datos. Además, los datos personales sensibles son datos personales relacionados con ámbitos como la salud, la política, la religión o la orientación sexual, que no pueden recopilarse ni tratarse más que en situaciones bien definidas. Estas nociones y las obligaciones asociadas constituyen la piedra angular de la normativa francesa y europea (por ejemplo, a través del Reglamento General de Protección de Datos europeo, RGPD), como se analizará en este capítulo. La noción de Información Personalmente Identificable (IPI), ampliamente utilizada en EE.UU.[86], se aproxima a la de datos personales tal y como se ha definido anteriormente, pero no es equivalente.

Las consideraciones de privacidad se han convertido en un tema central en nuestro mundo conectado. Varios ámbitos, aún no afectados por esta tendencia, pronto generarán enormes cantidades de datos personales y a veces sensibles, sin dejar a los usuarios ninguna opción de exclusión. La privacidad es, por tanto, una cuestión clave, tan importante como la seguridad.

Al ser lógicamente multifacético, el trabajo sobre la privacidad abarca varias dimensiones:

- una parte es jurídica: se necesitan normas armonizadas, que se apliquen en el mayor ámbito geográfico posible, para favorecer las buenas prácticas y prohibir las otras. Para que sea aplicable, puede ser necesaria la utilización de directrices para ejecutar estas normas armonizadas, y esto puede ser un verdadero reto;
- una parte es técnica: se necesitan herramientas avanzadas de privacidad tanto en el ámbito teórico como en el aplicado. Pueden ayudar a analizar y mejorar los sistemas existentes o a diseñar sistemas de preservación de la privacidad desde cero;
- una parte es económica: entender el ecosistema subyacente es esencial, ya que suele determinar las prácticas de recogida y tratamiento de datos personales. Es necesario un ecosistema sostenible que respete la normativa europea en materia de protección de datos;
- una parte es cultural: los pueblos de distintas zonas geográficas pueden

86. http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

tener enfoques diferentes sobre la privacidad, debido a sus raíces culturales, y estas diferencias repercuten en la normativa local en materia de datos personales;

- por último, una parte es sociológica: el usuario final suele declararse preocupado por la privacidad y, al mismo tiempo, se comporta de forma contraria.

Esta conocida "paradoja de la privacidad" pone de manifiesto la necesidad de realizar estudios sociológicos para comprender mejor los comportamientos humanos en este ámbito y mejorar potencialmente la concienciación y las prácticas.

Este capítulo cubre esencialmente los aspectos técnicos y las dimensiones legales hasta cierto punto, ya que los demás aspectos están fuera del ámbito de investigación de Inria. En primer lugar, presentamos los principios y la normativa de primer nivel en materia de privacidad. A continuación, presentamos las herramientas y las tecnologías de mejora de la privacidad. Por último, analizamos las filtraciones de privacidad en los sistemas existentes, donde los datos personales son comunicados deliberadamente por el usuario o recogidos sin su conocimiento.

5.1 Principios de privacidad y normativa

[Resumen]

Para tener en cuenta los grandes cambios que se produjeron durante la última década en materia de recopilación y uso de datos personales, la Unión Europea adoptó el Reglamento General de Protección de Datos ("RGPD" por sus siglas en inglés) que entró en efecto en mayo de 2018. El principal cambio es el énfasis puesto bajo la responsabilidad de los controladores de datos, es decir, las organizaciones que procesan datos personales, así como sus subcontratistas, si los hay. Todo responsable del tratamiento debe realizar evaluaciones de impacto sobre la protección de datos, aplicar la privacidad desde la fase de diseño y rendir cuentas. Si la evaluación de impacto indica que el tratamiento puede afectar gravemente a los derechos y libertades de las personas físicas, habrá que reforzar las medidas adoptadas. Los derechos del interesado también se ven fortalecidos con una mejor información y control sobre sus datos, siguiendo la filosofía de empoderamiento del usuario.

Sin embargo, el RGPD ofrece muy poca orientación sobre la aplicación eficaz de estos conceptos. Es necesario un trabajo interdisciplinario para reducir esta brecha entre los instrumentos jurídicos y técnicos, por ejemplo, por medio de la definición de un análisis riguroso del riesgo para la privacidad y métodos de privacidad por diseño, o estableciendo técnicas para reforzar la responsabilidad, la transparencia y mejorar el control del usuario sobre sus datos personales. En cualquier caso, la privacidad tiene un precio, ya que existen conflictos con otras consideraciones y la privacidad se considera a veces un factor limitante.

Un cierto número de principios fundamentales y consideraciones jurídicas rigen la privacidad. Esta sección los analiza, en particular los asociados al nuevo reglamento europeo (RGPD) y a otros textos legales (por ejemplo, el reglamento de ePrivacy que particulariza y complementa el RGPD).

5.1.1 Conflictos entre la privacidad y otras consideraciones

La privacidad tiene un precio. En un contexto en el que muchos servicios comerciales dependen de los datos personales (por ejemplo, varios servicios gratuitos se sustentan con publicidad personalizada), donde el big data puede ofrecer servicios de gran valor (por ejemplo, el estudio de la propagación del virus de la gripe a través del análisis de los documentos médicos), donde varios países despliegan sistemas de vigilancia masiva destinados a ayudar a combatir el terrorismo y donde se necesita una cierta forma de trazabilidad de los usuarios para permitir relaciones respetuosas entre los ciudadanos, la privacidad puede considerarse un factor limitante. Esta es la señal de una tensión fundamental entre la privacidad, que requiere minimizar la recogida de datos personales, y otras consideraciones como la utilidad, la seguridad o la responsabilidad, donde cuanto mayor sea el volumen y la precisión de los datos, mejor. Por lo tanto, es necesario llegar a un compromiso y la idea de encontrar un equilibrio adecuado es fundamental, por ejemplo, para la normativa de protección de datos. Este equilibrio, por supuesto, depende en gran medida de los aspectos culturales, de ahí la importancia de contar con una normativa europea en la materia para preservar nuestra soberanía. Las siguientes secciones abordarán e ilustrarán esta tensión según varios ángulos.

5.1.2 Evolución del marco normativo

La noción de privacidad es compleja y multifacética. Además, su percepción evoluciona a través del tiempo y el espacio y se ve afectada por la adopción de nuevas tecnologías. Para tener en cuenta los grandes cambios que se han producido durante la última década en cuanto a la recogida y el uso de datos personales, la Unión Europea adoptó en 2016 el "Reglamento General de Protección de Datos" (o RGPD) que entró en vigencia en mayo de 2018, en todos los Estados miembros de manera uniforme.

El mayor cambio introducido por el RGPD es el énfasis puesto en la responsabilidad del administrador del tratamiento de datos (es decir, la organización o asociación privada o pública que procesa datos personales), así como de su subcontratista (si lo hay). Todo responsable del tratamiento de datos debe:

- realizar evaluaciones del impacto de la protección de datos;
- implementar la privacidad por definición;
- y cumplir con el principio de rendimiento de cuentas.

Si la evaluación de impacto indica que el tratamiento puede afectar gravemente a los derechos y libertades de las personas físicas, habrá que reforzar las medidas adoptadas. También se refuerzan los derechos de los interesados para mejorar su información y control sobre sus datos personales, siguiendo un enfoque de empoderamiento del usuario. Por ejemplo, el responsable del tratamiento debe estar en condiciones de demostrar que ha obtenido el consentimiento explícito del usuario, los usuarios tienen el "derecho al olvido" en los servicios de búsqueda prestados en Europa, se añade un nuevo derecho de portabilidad de los datos para que un usuario pueda cambiar de plataforma reutilizando sus datos, y se protege mejor a los menores de 16 años.

Sin embargo, el RGPD ofrece muy poca orientación sobre la aplicación eficaz de estas nuevas disposiciones y algunas de ellas plantean una serie de problemas técnicos. Antes de describir la investigación sobre las tecnologías de mejora de la privacidad en el apartado 5.2, se ofrece una visión general de los retos que plantean los requisitos legales y los trabajos interdisciplinarios para reducir la brecha entre los instrumentos legales y los técnicos.

5.1.3 Evaluación de impacto relativa a la protección de datos (EIPD)

Las Evaluaciones de impacto relativas a la protección de datos (EIPD), o simplemente Evaluaciones de impacto sobre la privacidad (EIP), son utilizadas por las organizaciones para evaluar cualquier problema de privacidad que pueda surgir al desarrollar nuevos productos o servicios que impliquen el tratamiento de datos personales. La realización de una EIPD es obligatoria en Europa en virtud del RGPD para determinadas categorías de tratamiento de datos personales. Más allá de los requisitos legales, la realización de una EIPD redunda en el interés de cualquier organización para garantizar que los riesgos para la privacidad se comprendan y aborden adecuadamente antes de desplegar cualquier nuevo producto o servicio. Ya existe un amplio corpus de contribuciones sobre las EIPD y se han publicado varias EIPD para productos específicos. La Comisión Nacional de Informática y Libertades (CNIL), la agencia francesa de protección de datos, también ha publicado recientemente una herramienta para ayudar a los controladores de datos a preparar una EIPD.[87]

Todas estas contribuciones son muy útiles para definir el proceso de la EIPD (incluida la planificación, la consulta a las partes interesadas, la asignación de recursos o las auditorías) y su objetivo principal (evaluar la probabilidad y la gravedad de las amenazas a la privacidad). Sin embargo, no definen con mucha precisión cómo debe realizarse la parte técnica de la EIPD, el análisis del riesgo para la privacidad.

Por supuesto, hay una serie de puntos en común entre los análisis de riesgos

87. https://www.cnil.fr/fr/PIA-privacy-impact-assessment

de privacidad y de seguridad. Sin embargo, la privacidad es un concepto más complejo y multifacético que tiene como objetivo la protección de las personas (es decir, los individuos, los grupos y la sociedad en su conjunto) en lugar de los recursos o las organizaciones. Estas dimensiones deben tenerse en cuenta en un análisis de riesgo para la privacidad, especialmente la noción de daño a la privacidad que ha sido ampliamente discutida por los juristas.

Aunque se están definiendo marcos de análisis de riesgos para la privacidad, aún queda trabajo por hacer. En primer lugar, sería muy útil en la práctica disponer de marcos de análisis de riesgos para la privacidad específicos para determinados ámbitos de aplicación. Desde el punto de vista teórico, también sería interesante establecer vínculos formales con las métricas de privacidad. Por último, pero no por ello menos importante, más allá del cumplimiento legal, un gran beneficio de un análisis de riesgos para la privacidad debería ser proporcionar orientación para el diseño de un nuevo producto, siguiendo el enfoque de privacidad por definición.

5.1.4 Privacidad por definición (PbD)

La filosofía de la privacidad desde el diseño, que es obligatoria en virtud del RGPD, es que ésta no debe tratarse como una idea secundaria, sino como un requisito de primera clase en el diseño de cualquier sistema. Sin embargo, desde un punto de vista técnico, sigue siendo un reto:

- la privacidad incluye una serie de dimensiones (como la limitación de la recogida, la calidad de los datos, la especificación de su finalidad, la limitación del uso o la seguridad) que generalmente no se definen con mucha precisión;
- entonces, estos requisitos pueden parecer estar en conflicto con otros requisitos como los funcionales, la facilidad de uso, el rendimiento o la viabilidad económica del producto o servicio.

Para implementar la privacidad por diseño, existe una amplia gama de Tecnologías que potencian la privacidad (PET, o Private Enhacing Technologies en su acepción inglesa), como se explica a continuación. Cada PET ofrece distintas garantías basadas en distintos supuestos y, por tanto, es adecuada en distintos contextos. En consecuencia, es bastante complejo para un ingeniero de software elegir con conocimiento de causa entre todas estas posibilidades y encontrar la combinación de técnicas más adecuada para resolver sus propios requisitos. Se han propuesto soluciones en distintos ámbitos de aplicación, como la medición inteligente, el pago por uso o los sistemas basados en la localización, pero el siguiente reto en este ámbito es ir más allá de los casos individuales y establecer bases y metodologías sólidas para la privacidad por diseño.

Un marco formal, por ejemplo basado en la lógica epistémica, puede ser útil para expresar los requisitos de minimización de datos como propiedades que definen para cada parte interesada la información que se le permite conocer o no. Sin embargo, los requisitos en conflicto suelen tener que cumplirse

simultáneamente, por ejemplo, las garantías sobre la corrección del resultado de un cálculo. En efecto, la pugna entre la minimización de los datos y la corrección es una de las cuestiones delicadas que hay que resolver en muchos sistemas que implican datos personales. Un marco formal también es útil para razonar sobre las arquitecturas. Su axiomatización puede utilizarse para demostrar que una determinada arquitectura cumple con los requisitos de privacidad e integridad previstos.

5.1.5 Rendición de cuentas

En consonancia con normativas anteriores (por ejemplo, las Directrices de la OCDE de 1980 sobre la protección de la privacidad y los intercambios transfronterizos de datos personales), el RGPD integra la responsabilidad como un principio que exige que las organizaciones pongan en marcha las medidas técnicas y organizativas adecuadas para demostrar su cumplimiento del reglamento.

Se puede distinguir entre tres tipos principales de responsabilidad:

- La responsabilidad de la política puede considerarse el primer nivel de rendición de cuentas: la organización debe ser capaz de demostrar que ha definido una política de privacidad clara y debidamente documentada;
- La responsabilidad de los procedimientos, que se refiere a la prueba de los mecanismos organizativos, como los procesos documentados para hacer frente al consentimiento de los usuarios o a las reclamaciones o solicitudes de datos personales;
- La responsabilidad de la práctica es la demostración a posteriori de la efectividad de la responsabilidad aplicada a los procedimientos. Es una prueba de que las políticas de privacidad se han cumplido efectivamente.

El primer tipo de rendición de cuentas es puramente declarativo y proporciona, en el mejor de los casos, una forma de garantía legal (compromiso vinculante). El segundo tipo añade garantías a nivel organizativo, pero sólo el tercer tipo puede cumplir todas las expectativas de responsabilidad. Sin embargo, hasta ahora se ha hecho hincapié en el primer y segundo tipo de rendición de cuentas, lo que ha dado lugar a garantías superficiales.

Para que la rendición de cuentas contribuya efectivamente a la mejora de la protección de la privacidad, es necesario poder traducir sus principios generales en medidas prácticas y considerar sus diversas dimensiones. Por ejemplo, la recopilación y el mantenimiento de registros de tratamiento detallados pueden ser contradictorios con la privacidad si esos registros contienen datos personales. Hay que encontrar un equilibrio adecuado entre estos requisitos.

5.1.6 Empoderamiento del usuario mediante el control y la transparencia

Un servicio que cumple con la legislación y del que es responsable el encargado (o responsable) del tratamiento de datos no está necesariamente en consonancia

con a las expectativas del usuario. Es necesario empoderar al usuario para conseguir su aceptación.

EL EMPODERAMIENTO MEDIANTE EL CONTROL DEL USUARIO

La privacidad se considera cada vez más como la capacidad del usuario de controlar sus datos personales. Sin embargo, si bien esta noción predomina en la literatura sobre privacidad y desempeña un papel central en el RGPD, siguen faltando definiciones claras. La palabra "control" suele utilizarse de forma muy vaga en este contexto, tanto por los juristas como por los informáticos. Por ejemplo, nociones como "control de acceso" o "control de uso" en informática no incorporan realmente la intuición subyacente a la noción de control sobre los datos personales. Esta falta de precisión puede dar lugar a malentendidos y dificulta la comprobación de su cumplimiento. Un estudio multidisciplinar de la noción de control tal y como la utilizan los juristas y los informáticos ha llevado a identificar tres dimensiones, que se corresponden con las capacidades de un individuo:

- para realizar acciones sobre sus datos personales;
- para evitar que otros realicen acciones sobre sus datos personales;
- ser informado de las acciones realizadas por otros sobre sus datos personales.

En la práctica, deben cumplirse dos condiciones principales para que los interesados puedan ejercer el control sobre sus datos personales:

- el usuario debe estar debidamente informado sobre la recopilación de sus datos, su finalidad, la entidad que los recoge y, por ejemplo, el plazo de conservación (transparencia);
- el usuario debe poder expresar su elección de que se recopilen o no sus datos para un fin determinado y tener garantías de que esa elección se cumple realmente (consentimiento).

En efecto, el consentimiento explícito del interesado, que es una piedra angular de la mayoría de las normas de protección de datos, es un ejemplo típico de un requisito legal que es muy difícil de poner en práctica. Este es el caso de la Internet de las cosas (IoT) donde muchas comunicaciones de datos se producen sin que el usuario lo advierta, o las redes sociales, por ejemplo, y se han hecho varias propuestas para ayudar al usuario a ejercer su control.

EL EMPODERAMIENTO A TRAVÉS DE LA TRANSPARENCIA

La transparencia es un concepto esencial para diseñar sistemas y servicios que preserven la privacidad. El responsable del tratamiento de datos debe proporcionar información clara y completa sobre qué información se recoge, con qué recurrencia, para qué, cómo se procesan los datos, cómo se almacenan los datos (dónde, cuánto tiempo, con qué seguridad), y si es probable que los datos

se comuniquen a terceros. Las mismas preguntas se aplican recursivamente a los terceros a los que se pueden transferir los datos.

La transparencia es un reto en diversos ámbitos en los que la recogida de datos se produce de forma invisible. Este puede ser el caso de los sistemas de la Internet de las Cosas (IoT) que miden y recogen continuamente datos personales. La información proporcionada a los sujetos de los datos debe ser lo más visible e inteligible posible (lo que excluye los simples carteles en las paredes que generalmente pasan desapercibidos). Se requiere un nuevo espacio de diseño para los avisos de privacidad efectivos, que es un tema de investigación activo en el área de la privacidad.

Más allá de la transparencia en la recogida de datos, un nuevo reto importante es la transparencia de los algoritmos. Esta necesidad de mayor transparencia es otra ilustración de una exigencia legal (en el RGPD y también en la ley para una República digital[88] adoptada en Francia en octubre de 2016) que plantea muchos retos técnicos y puede ser una fuente de temas de investigación interdisciplinar.

La transparencia de los algoritmos es fundamental en muchos sistemas automatizados con impacto en la sociedad (por ejemplo, con un sistema automatizado de asignación, una vez que todos los candidatos han publicado sus preferencias solicitadas): la cuestión de un posible sesgo algorítmico, ya sea intencionado o no (por ejemplo, un error), es inevitable. La transparencia de los algoritmos es necesaria para que terceros (incluidos los ciudadanos) puedan analizar su comportamiento interno. Este es el papel de la iniciativa francesa TransAlgo[89], dirigida por Inria.

Dado que la transparencia requiere que se comunique al usuario información detallada, la forma de conseguirla no es trivial. El enfoque habitual, a través de las Condiciones Generales de Uso, no suele ser satisfactorio, ya que no es fácil para el usuario: este documento legal suele estar destinado a proteger a la empresa más que a informar al usuario. Se está investigando sobre el análisis de las prácticas, sobre el intento de tener un formato estandarizado comprensible para todo el mundo (como hizo Creative Commons en un dominio diferente) y quizás susceptible de tratamiento mecanizado, y sobre los impactos de la transparencia en el comportamiento de los usuarios.

[Desafío de investigación 7]Entender la privacidad y obtener herramientas prácticas

Entender los principios y la normativa sobre privacidad es la base de cualquier actividad en este ámbito. Aunque no se trata de un ámbito nuevo (por ejemplo, la "Loi Informatique et Libertés" se adoptó en 1978), esta área ha experimentado recientemente importantes evoluciones con el nuevo reglamento europeo RGPD y, al mismo tiempo, nuevas oportunidades de recopilar datos personales. En consecuencia, la

88. Loi pour une Republique numerique. https://www.economie.gouv.fr/republique-numerique.
89. https://www.inria.fr/en/news/news-from-inria/transalgo

comprensión de los conceptos y de la normativa es una primera necesidad. Otra es la de poder obtener herramientas prácticas: aunque el RGPD promueve varios conceptos y objetivos, ofrece poca orientación sobre la aplicación eficaz de estas nuevas disposiciones reglamentarias.

En particular, el RGPD introdujo el derecho a la portabilidad de los datos, según el cual un usuario puede recuperar sus datos en un formato legible por el ser humano y por la máquina. Este derecho abre nuevos campos de investigación en torno a la gestión y el control individualizados de los datos personales. El objetivo es capacitar a los ciudadanos para que aprovechen sus datos personales para su propio bien, lo que exige plataformas personales en la nube seguras, extensibles y soberanas, tres objetivos en conflicto que abren nuevos retos de investigación (véase, por ejemplo, el apartado 5.2.4).

[Equipos Inria] Principios de privacidad y normativa

↗ El equipo **CIDRE** trabaja en políticas de privacidad y en el derecho al olvido en colaboración con abogados.

↗ El equipo **INDES** trabaja en las relaciones entre el rastreo en la web y el Reglamento de privacidad electrónica, en relación con abogados. En particular, el equipo evalúa las repercusiones sobre la privacidad cuando el Reglamento no exige el consentimiento del usuario para el rastreo y crea herramientas que detectan las violaciones del Reglamento.

↗ El equipo **PETRUS** trabaja en el control de los datos personales y la minimización de los mismos, y mantiene fuertes colaboraciones con grupos de investigación de otras disciplinas como la economía, el derecho y las ciencias sociales.

El equipo también contribuye a mejorar el control de los individuos sobre sus datos personales desde un punto de vista de la arquitectura. Este es el caso de PlugDB, un servidor personal seguro que permite a los individuos ejercer el control sobre sus datos personales, preservando la durabilidad, la disponibilidad y el intercambio.

↗ El equipo **PRIVATICS** trabaja en la mayoría de los temas, con un fuerte énfasis en la interdisciplinariedad a través de colaboraciones con abogados y economistas. Por ejemplo, el equipo ha contribuido a crear un marco y una metodología para llevar a cabo un análisis del riesgo para la privacidad de forma rigurosa y sistemática, compatible con la mayoría de las recomendaciones de la EIPD. En cuanto a la privacidad por diseño, el equipo ha propuesto un marco formal basado en la lógica epistémica que permite expresar los requisitos de minimización de datos. Este marco se ha utilizado para comparar formalmente distintas arquitecturas de control de acceso biométrico. En cuanto a la responsabilidad, el equipo ha definido un conjunto de medidas prácticas que deben adoptarse en cada fase del ciclo de vida de los datos personales, desde la recogida hasta la eliminación, pasando por el almacenamiento, el uso y la transmisión a terceros. Desde el punto de vista formal, el equipo ha propuesto un marco basado en registros que respetan la privacidad, demostrando que el cumplimiento puede

comprobarse a partir de registros que no contienen datos personales.
En cuanto al control del usuario, tras un estudio multidisciplinar de esta noción tal y como la utilizan los juristas y los informáticos, se ha derivado un modelo formal para caracterizar formalmente cada tipo de control.
En el caso particular de la IoT y el control de los usuarios, el equipo ha propuesto una arquitectura basada en "Agentes de Privacidad" que implementan las elecciones del sujeto de los datos, expresadas en un lenguaje natural restringido que puede ser fácilmente entendido por los no expertos.

5.2 Herramientas de privacidad

[Resumen]
Para pasar de los principios fundamentales del RGPD a los productos y servicios conformes con la privacidad, se necesitan herramientas para distintos tipos de público, desde el responsable de la protección de datos hasta el usuario final. Algunos de ellos se centran en principios tales como la evaluación del impacto de la protección de datos, la privacidad por definición y la responsabilidad.
Otras herramientas están pensadas para anonimizar una base de datos antes de liberarla para el acceso a datos abiertos. Sin embargo, se trata de una tarea compleja que requiere encontrar un equilibrio adecuado: aumentar la privacidad suele reducir la utilidad de una base de datos anonimizada. Desde este punto de vista, el concepto de privacidad diferencial ha resultado ser una herramienta clave para proporcionar garantías de privacidad demostrables.
La llegada de las nubes personales, destinadas a dar a los usuarios un control total sobre sus propios datos, es otra de las principales herramientas de empoderamiento del usuario. Por último, se han diseñado técnicas para proporcionar la desvinculación, es decir, la garantía de que nadie puede vincular varios usos de un servicio por un usuario determinado. Esto es especialmente importante para los sistemas que incorporan un token RFID, como los pasaportes electrónicos. En cuanto a las comunicaciones, el sistema Tor intenta garantizar las comunicaciones anónimas (nadie puede identificar el origen de un paquete), aunque en la práctica esto es menos utilizable.

5.2.1 Herramientas relacionadas con la EIPD, la privacidad por diseño y la responsabilidad

Se han diseñado varias herramientas para responder a las necesidades de evaluación del impacto de la protección de datos, privacidad por diseño y responsabilidad. Algunas de ellas se analizan respectivamente en los apartados 5.1.3, 5.1.4 y 5.1.5.

5.2.2 Anonimización de bases de datos: una necesidad para los datos abiertos y el big data

Todo responsable de una base de datos que contenga datos personales está sujeto a una normativa estricta. Es posible eludir esta normativa anonimizando esta base de datos, mientras la base de datos resultante ya no contenga ningún dato personal. Por ejemplo, este es un requisito previo para liberar un conjunto de datos públicos en el contexto de una iniciativa de datos abiertos.

La anonimización de datos consiste en alterar el conjunto de los mismos para eliminar cualquier información que pueda utilizarse para volver a identificar a cualquier participante del conjunto de datos o para inferir atributos personales. No es una tarea fácil. Una primera razón es que la anonimización es intrínsecamente compleja y depende del ámbito: no existe una solución universal. Algunos tipos de datos, como los rastros de movilidad, son muy exclusivos y, por tanto, identificables. Por ejemplo, el conocimiento de cuatro puntos espacio-temporales puede ser suficiente para identificar de forma inequívoca al 95% de los individuos de una gran base de datos de movilidad de un operador de telefonía móvil[90][dMHVB13].

Una segunda razón es que el anonimizar una base de datos es una cosa y evitar la re-identificación a través de información secundaria (el problema de la "inferencia") es otra. Desde este punto de vista, es famosa la desanonimización del seudónimo nº 4417749 en el caso de AOL[91][BZ06]. Al realizar un análisis de los registros detallados de las búsquedas que éste hizo, se encontró rápidamente su identidad. En general, se sabe que la sustitución de nombres por seudónimos es muy vulnerable a los ataques de re-identificación.

Por último, pero no por ello menos importante, es necesario un compromiso adecuado entre privacidad y utilidad (véase el apartado 5.1.1). Garantizar una privacidad significativa requiere la distorsión del conjunto de datos original, que mecánicamente produce un conocimiento impreciso y de grano grueso incluso sobre la población en su conjunto.

La anonimización de datos puede lograrse mediante varios tipos de enfoques. En 2014, el grupo G29, que reúne a las agencias europeas de protección de datos, publicó un documento técnico sobre el tema[92][Art14]. Este documento analiza la eficacia y los límites de varias técnicas: permutación, privacidad diferencial (véase el apartado 5.2.3), agregación, k-anonimato, l-diversidad y t-cercanía. En términos de garantías de privacidad, estas técnicas pueden clasificarse en dos categorías: modelos de privacidad sintácticos y semánticos. Los modelos sintácticos se centran en los requisitos sintácticos de los datos anonimizados (por ejemplo, con el anonimato k cada registro debe aparecer al menos k veces en el conjunto de datos anonimizados),

90 [dMHVB13] Yves-Alexandre de Montjoye, Cesar A. Hidalgo, Michel Verleysen, y Vincent D. Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3:1376, 03 2013.
91 [BZ06] M. Barbaro y T. Zeller. A face is exposed for aol searcher no 4417749. New York Times, agosto de 2006.
92 [Art14] Dictamen 05/2014 sobre técnicas de anonimización, abril de 2014.

sin ninguna garantía sobre la información sensible que un atacante puede conocer exactamente sobre los individuos. Los modelos semánticos, en cambio, se ocupan de la información privada que puede deducirse sobre los individuos utilizando los datos anonimizados, así como quizás algún conocimiento previo (o de fondo) sobre ellos. De los modelos semánticos cabe esperar un mayor nivel de privacidad. Este es el caso de la privacidad diferencial que se discute en la siguiente sección.

5.2.3 Privacidad diferencial

La privacidad diferencial[93][DMNS06] se propuso originalmente en el ámbito de las bases de datos estadísticas y hoy en día es uno de los enfoques más exitosos de la privacidad. El objetivo es proteger los datos de un individuo al tiempo que se publica información agregada sobre la base de datos. Esto se consigue añadiendo ruido controlado al resultado de la consulta, de manera que los datos de un solo individuo tengan un impacto insignificante en la respuesta notificada. La privacidad diferencial tiene varias ventajas: (1) es independiente de la información colateral del adversario, lo que significa que no es necesario tener en cuenta el contexto en el que operará el sistema; (2) es compositivo, es decir, si combinamos la información que obtenemos al consultar dos mecanismos diferencialmente privados, el mecanismo resultante también es diferencialmente privado; y (3) los mecanismos diferencialmente privados suelen ofrecer un buen equilibrio entre utilidad y privacidad.

Una variante exitosa, llamada privacidad diferencial local, tiene la ventaja de no requerir un tercero de confianza: los usuarios encubren sus datos personales, añadiendo ruido por sí mismos, antes de enviarlos al recopilador de datos. La privacidad diferencial local es especialmente adecuada cuando los datos se recogen con fines estadísticos, ya que suele lograr un buen equilibrio entre privacidad y utilidad. La privacidad diferencial local tiene las mismas ventajas que la privacidad diferencial (independencia del conocimiento lateral y composicionalidad). Tuvo un impacto considerable después de que grandes empresas como Apple y Google la adoptaran para los sistemas de recogida de datos que preservan la privacidad (por ejemplo, Google utiliza una implementación particular en la tecnología de crowdsourcing RAPPOR).

Existen muchos casos en los que el dominio de los datos presenta una noción de distancia (por ejemplo, la ubicación, el consumo de energía en los contadores inteligentes, o la edad y el peso en los registros médicos). En ese caso, la relación privacidad/utilidad puede mejorarse en gran medida explotando el concepto de aproximación intrínseco a la noción de distancia. La idea es permitir que dos

93 [DMNS06] C. Dwork, F. McSherry, K. Nissim, y A. Smith. Calibrating noise to sensitivity in private data analysis. En Theory of Cryptography (TCC'06), 2006. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

valores sean cada vez más distinguibles a medida que su distancia aumenta para poder realizar estadísticas más precisas.

[Desafío de investigación 8] Datos abiertos y anonimización

Las iniciativas de datos abiertos pueden significar a veces la publicación de bases de datos que contienen información personal sensible. Para garantizar la privacidad de los individuos, los datos deben ser anonimizados. La anonimización robusta, que resiste eficazmente los ataques de desanonimización, es un tema de investigación activo y candente. Aunque la privacidad diferencial se ha convertido en una herramienta científica clave para lograr garantías de anonimización demostrables, siguen existiendo retos en su aplicación, por ejemplo, para mejorar la relación entre privacidad y utilidad.

5.2.4 El empoderamiento de los usuarios mediante nubes personales

Ha llegado el momento de la gestión y el control individualizado de los datos personales. Gracias a las iniciativas de divulgación inteligente y al derecho a la portabilidad de datos del RGPD, podemos acceder a nuestros datos personales desde las empresas u organismos gubernamentales que los recogieron, lo que permite la creación de repositorios de datos personales entre dominios. Al mismo tiempo, las soluciones de Nube Personal, también llamadas Sistema de Gestión de Información Personal (PIMS, o Personal Information Management System) o Servidor de Datos Personales (Personal Data Server por su denominación inglesa, PDS), están proliferando (véase, por ejemplo, la empresa francesa Cozy Cloud[94]). Su objetivo es permitirnos aprovechar nuestros datos personales para nuestro propio bien, abriendo el camino a nuevos servicios de valor añadido mediante el cruce de datos emitidos desde distintos silos de datos o compartiendo nuestros datos para obtener beneficios sociales/comunitarios (por ejemplo, contribuir a un estudio epidemiológico, calcular consultas basadas en datos compartidos dentro de comunidades de usuarios).

Sin embargo, la gestión de nuestros propios datos personales constituye una carga considerable. Debemos garantizar la seguridad de los datos que recopilamos y gestionar los datos divulgados y controlar su uso. Heredamos la responsabilidad combinada de un experto en seguridad de la información y un administrador de bases de datos. Por ello, los proveedores de nubes personales proponen soluciones para gestionar los datos personales en nombre de sus clientes, creando así auténticos puntos neurálgicos (centralizando enormes cantidades de datos personales pertenecientes a millones de individuos). Así, paradójicamente, al empoderar a los usuarios, la divulgación inteligente y las nubes personales los empujan hacia una delegación aún mayor sobre más datos, exponiéndolos así a un riesgo mayor que nunca.

94. https://cozy.io/en/

Prototipo de nube personal específico – © Inria / Photo C. Morel

Para escapar de esta paradoja y lograr realmente el empoderamiento del usuario, se deben cumplir una serie de exigencias: (1) soberanía: el sistema debe ofrecer al usuario la capacidad de ejercer sus decisiones de divulgación de datos bajo su propia autoridad y sin ninguna forma de delegación; (2) seguridad: el sistema debe ofrecer garantías tangibles sobre el cumplimiento de estas decisiones sea cual sea el tipo de ataques o usos indebidos a los que pueda enfrentarse el sistema, no sólo al titular de la nube personal, sino también a los demás usuarios y a terceros; y (3) extensibilidad: el sistema no debe impedir el desarrollo de nuevos servicios que utilicen los datos de un solo individuo ("Personal Big Data") o de grandes grupos de individuos ("Big Personal Data"). Sin embargo, estas afirmaciones introducen dos formas de incertidumbre, una primera entre la soberanía y la seguridad: conciliar la ausencia de delegación en expertos o administradores centrales de IT con altas garantías de seguridad; y otra entre la seguridad y la extensibilidad: conciliar la seguridad, que exige sistemas cerrados, con la necesidad de extensibilidad para soportar nuevos servicios de datos, que por definición no son de plena confianza. Conseguir una arquitectura que cumpla con estos tres requisitos es una cuestión intrínsecamente difícil.

Otro problema importante relacionado con el empoderamiento de los usuarios es permitir nuevas aplicaciones de grandes datos personales (por ejemplo, detección participativa, estudios epidemiológicos y sistemas de recomendación personalizados). El procesamiento disperso de datos personales no es un tema nuevo. Sin embargo, la nube personal está distribuida a nivel individual y se

espera que se amplíe a poblaciones de todo el país. En este contexto, la primera cuestión es evitar o minimizar el efecto de las fugas de datos personales durante los cálculos realizados a tal escala. Además, la descentralización hace hincapié en el papel central del individuo en la arquitectura. Esto exige la aparición de nuevas formas de computación descentralizada en las que los perfiles individuales y las configuraciones de privacidad individuales se integran por construcción.

Más allá de la perspectiva de la privacidad y la seguridad, quedan por resolver muchos problemas importantes de investigación en torno a las nubes personales: aspectos fundacionales y de sistemas de gestión de datos complejos, especialmente con datos centrados en el ser humano y orquestación de consultas a los distintos servicios.

5.2.5 Protocolos y tecnologías de comunicación que preservan la privacidad

PROTOCOLOS DE DESVINCULACIÓN

Según la norma ISO/IEC 15408-2, la desvinculación, también conocida como la no rastreabilidad, "garantiza que un usuario pueda hacer múltiples usos de recursos o servicios sin que otros puedan vincular estos usos". Por supuesto, esta noción de desconexión se aplica tanto en el mundo físico (es decir, el rastreo de una persona) como en el mundo virtual (es decir, la vinculación de las transacciones de una persona).

La desvinculación es cada vez más importante con el uso generalizado de tokens RFID para la autenticación. Si el protocolo de autenticación garantiza la desvinculación, debería ser imposible colocar un lector RFID no autorizado que pueda decidir si un token determinado es el mismo que otro visto anteriormente. Por ejemplo, dado que los tokens RFID se utilizan en los pasaportes electrónicos, un atacante que intercepte una primera ejecución legítima del protocolo para un objetivo, digamos Alice, no debería ser capaz de instalar un lector que distinga el pasaporte de Alice de otros. Del mismo modo, los protocolos de autenticación implementados en las redes de telefonía móvil 3G/4G/5G impiden que un espía, que no sea el operador, pueda relacionar comunicaciones diferentes realizadas por el mismo dispositivo.

Del mismo modo, las transacciones digitales pueden requerir la desvinculación. Un ejemplo típico es una moneda electrónica basada en una cadena de bloques (blockchain) en la que el libro de registro es completamente público. Este tipo de sistemas suelen proporcionar 'seudonimato' (es decir, el uso de un seudónimo en lugar de una identidad real). Sin embargo, al vincular varias transacciones realizadas por el mismo usuario, el perfilado del usuario es trivial y luego la re-identificación puede ser posible con información colateral. Dicho esto, Zerocash[95] es un ejemplo de moneda electrónica que ofrece garantías de desvinculación.

95. http://zerocash-project.org/

De forma más general, a veces es necesario autenticar y autorizar garantizando el anonimato y la no vinculación. En ese caso, los protocolos de autenticación anónima directa (DAA, Direct Anonymous Authentication) proporcionan tokens de autenticación anónima, permitiendo la autenticación anónima remota.

Por último, el fingerprinting es una amenaza directa para la desvinculación. La "huella digital" del navegador se trata en el apartado 5.3.4. Sin embargo, incluso los dispositivos pueden ser rastreados por huella digital: por ejemplo, la longitud del mensaje encriptado que contiene datos personales, en particular la foto JPEG del pasaporte electrónico, puede ser efectivamente utilizada para el rastreo[96].

COMUNICACIONES ANONIMIZADAS

Tor es un sistema de anonimización que pretende preservar el anonimato de los usuarios de Internet. Esto se consigue gracias al llamado "enrutamiento cebolla": en lugar de conexiones directas desde un host a un destino, se crea un circuito virtual a nivel de aplicación que pasa por un cierto número de nodos de retransmisión repartidos por todo el mundo y elegidos al azar. El tráfico es encriptado por el cliente tantas veces como repetidores haya en este circuito virtual. Luego, cada nodo descifra la parte exterior del paquete recibido, revelando así la dirección IP del siguiente repetidor, y reenvía el paquete a éste. De este modo, Tor evita que cualquier fisgón identifique los hosts de origen y destino mirando las direcciones IP de origen y destino del paquete. Sin embargo, aunque sea teóricamente interesante, esta solución es muy difícil para el público en general debido a la cantidad de conocimientos necesarios para operar correctamente sobre Tor. De hecho, varios tipos de ataques de desanonimización contra Tor se encuentran regularmente.

[Equipos Inria] Herramientas de privacidad

↗ El equipo **COMETE** trabaja en la privacidad diferencial y sus variantes, en particular cuando existe una noción de distancia, así como en técnicas para medir la utilidad del resultado. Por ejemplo, el equipo ha implementado la privacidad dX y trabaja en una herramienta para recuperar con la mayor precisión posible la distribución original a partir de la que presenta ruido. El equipo también desarrolla una herramienta para medir la utilidad del resultado en términos de la calidad de la aproximación de la distribución verdadera. Por último, el equipo también trabaja en la definición de la desvinculación y su verificación formal.

↗ El equipo **DIANA** ha trabajado en ataques de desanonimización en Tor.

↗ El equipo **PESTO** trabaja en la definición de la desvinculación y su verificación formal.

↗ El equipo **PETRUS** trabaja en soluciones y arquitecturas para sistemas personales en la nube y en entornos de confianza para la computación descentralizada que preserva la privacidad. En particular, el equipo trabaja en una nueva arquitectura de referencia en

96. https://www.inria.fr/centre/nancy/actualites/securite-des-donnees-les-passeports-biometriques

la que las manipulaciones potencialmente complejas de los datos personales dependen de entornos de ejecución confiable (TEE, Trusted Execution Environments). El objetivo es limitar los efectos secundarios en términos de fugas de datos, ya que solo se declaran los resultados esperados a terceros, sin ningún acceso directo a los datos en bruto.

↗ El equipo **PRIVATICS** trabaja en diversas herramientas para el análisis del riesgo de la privacidad, la privacidad por diseño, la responsabilidad y el control de los usuarios. El equipo también trabaja en diferentes aspectos y modelos de anonimización de datos. El equipo ha desarrollado varios esquemas de anonimización para rastros móviles y conjuntos de datos de series de valores bajo los modelos de privacidad tipo k-anonimato o diferencial. El equipo también trabajó en una técnica novedosa para la liberación privada de conjuntos de datos de alta dimensión utilizando redes neuronales generativas. La idea es producir un conjunto de datos sintético que se parezca lo más posible a los datos de entrenamiento originales y que, al mismo tiempo, cumpla con los requisitos de privacidad. La privacidad diferencial se ha aplicado para proteger la privacidad del usuario en varios dominios, como los medidores inteligentes, la publicación de datos secuenciales, las redes neuronales generativas y el almacenamiento de datos en filtros de Bloom.

↗ El equipo **VALDA** estudia tanto los aspectos básicos como los sistémicos de la gestión de datos complejos, especialmente los centrados en el ser humano, en el contexto de las nubes personales. Se centran en las nubes personales bajo el ángulo de la integración de datos y la organización de servicios.

↗ El equipo **WIDE** trabaja en el uso de protocolos de gossip para preservar la privacidad de los cálculos descentralizados y busca extender este trabajo a la protección de la privacidad de los sistemas de aprendizaje automático descentralizados.

5.3 Análisis de la privacidad de los sistemas existentes

[Resumen]

Nuestro mundo conectado está en el origen de muchas filtraciones de privacidad, y a medida que pasa el tiempo, los dominios que hoy están libres de filtraciones también se verán afectados. Algunas de las filtraciones son deliberadas. Es el caso de las redes sociales, en las que los usuarios comparten masivamente (a veces en exceso) datos personales. Las consecuencias son numerosas para la privacidad del usuario. Por ejemplo, la caracterización de perfiles permitió el surgimiento de empresas dedicadas a comercializar servicios destinados a influir en los usuarios (por ejemplo, sus votos).

La información de geolocalización es otro tipo de información que se comparte con el consentimiento tácito o informado del usuario. Sin embargo, el registro de la ubicación de un usuario a lo largo del tiempo es particular en el sentido de que se pueden inferir muchas cosas, desde la ubicación de su casa y su trabajo hasta información personal sensible (por ejemplo, su religión si acude regularmente a un lugar de culto). Para beneficiarse de los servicios geolocalizados sin filtrar demasiada información, se han diseñado varias soluciones. Por ejemplo, con la ofuscación espacial, se reduce la precisión de la posición al informar una zona. La geo-indistinguibilidad, que aprovecha la privacidad diferencial, es una técnica prometedora para lograr la indistinguibilidad espacial con buenas propiedades adicionales.

También se puede invitar al usuario a proporcionar características biométricas, como las huellas dactilares, para permitir su identificación y autenticación, en el caso de los sistemas de control de acceso o los documentos nacionales de identidad seguros. Sin embargo, estos rasgos altamente diferenciadores y constantes (no pueden cambiarse) crean importantes riesgos de seguridad y privacidad. Los sistemas biométricos deben diseñarse cuidadosamente y, para investigar las opciones de diseño, se han desarrollado varios marcos para definir las arquitecturas de privacidad, para razonar formalmente sobre ellas y para justificar las opciones de diseño en términos de hipótesis de confianza.

Pero muy a menudo, las fugas de privacidad se producen sin el consentimiento del usuario. Este es el caso de la navegación por la web. Cada visita a un sitio web puede desencadenar una gran variedad de intercambios de datos ocultos entre múltiples empresas de seguimiento. La información puede utilizarse entonces para la publicidad dirigida, pero también para discriminar a los usuarios (por ejemplo, mediante precios personalizados) o algo peor. Se han propuesto varias soluciones en materia de privacidad, desde normativas de protección (por ejemplo, con el RGPD) hasta mecanismos del lado del cliente como herramientas de bloqueo de anuncios y rastreadores. Sin embargo, el ámbito está en constante evolución, con la aparición de nuevas técnicas para proteger mejor a los usuarios y, simultáneamente, para mejorar el rastreo dentro de un navegador o entre dispositivos.

Con la llegada de los teléfonos inteligentes y la IoT, las filtraciones de privacidad han alcanzado un volumen y una precisión sin precedentes, tanto en el mundo digital como en el físico, y a menudo sin el conocimiento del usuario. Esta tendencia se reforzará en los próximos años, abarcando nuevos dominios. Los objetivos de la investigación en este ámbito son analizar estos sistemas, dar información transparente de los comportamientos ocultos, poner de relieve las buenas y malas prácticas, proponer métodos susceptibles de mejorar la transparencia y el control de los usuarios, y animar a determinadas partes interesadas a cambiar de prácticas.

Como denominador común de nuestro mundo conectado, Internet también puede ser fuente de fugas de privacidad ocultas. Un primer ejemplo es la red de acceso inalámbrico

utilizada por la mayoría de los dispositivos. Hemos visto el rápido crecimiento de los sistemas de rastreo cibernético que analizan las tramas wifi enviadas por un smartphone en busca de un punto de acceso conocido. Es necesario investigar para analizar estas tecnologías y proponer versiones que preserven la privacidad siempre que sea posible. Otro ejemplo es el servicio DNS utilizado para asignar nombres de host a direcciones IP. Un atacante motivado podría espiar el tráfico del DNS (bajo ciertas circunstancias) y potencialmente identificar sitios web de interés. Desgraciadamente, todavía queda mucho camino antes de que se estandarice y despliegue una versión del DNS que respete la privacidad.
Por último, las técnicas de navegación segura utilizadas para detectar las URL de una lista negra (por ejemplo, las que se sabe que contienen un malware) son potencialmente intrusivas: almacenar toda la relación de URLs de la lista negra en la computadora es imposible y depender de un servicio externo es demasiado intrusivo, ya que este último recogería todas las URL visitadas. Se utilizan soluciones intermedias que deben ser analizadas cuidadosamente desde el punto de vista de la privacidad.

Esta sección se centra en las fugas de privacidad en varios sistemas existentes. Comienza con las filtraciones para las que existe un consentimiento tácito o informado del usuario y luego presenta las recopilaciones de datos que tienen lugar sin el conocimiento del usuario.

5.3.1 El lado visible: el caso de las redes sociales

Las redes sociales son lugares donde se incita a los usuarios a compartir datos personales de forma masiva. Muy a menudo, la información personal, e incluso la información personal sensible, se comparte con un grupo de personas que es notablemente mayor de lo que el usuario podría pensar. Varias razones pueden explicar esta situación. En primer lugar, la "paradoja de la privacidad" es un fenómeno bien conocido por el que los usuarios explican que están preocupados por su privacidad en línea, pero al mismo tiempo se comportan de forma contraria. Es posible que muchos usuarios no se den cuenta de que los riesgos les conciernen a ellos, sobre todo a los más jóvenes, para quienes los riesgos son teóricos o solo para los que "tienen algo que ocultar". O tal vez consideren que la recompensa obtenida al compartir información personal es superior a los riesgos. En segundo lugar, los usuarios pueden ser demasiado confiados en la protección ofrecida por la redes sociales, obviando el pequeño detalle de que éstas amplían significativamente el grupo de personas con las que se comparte la información. En tercer lugar, la configuración y los comportamientos por defecto con respecto a terceros también pueden ser mucho más permisivos de lo que cabría esperar (por ejemplo, Facebook cambió la audiencia por defecto a "Solo amigos" en 2014;

antes de ello las publicaciones eran "Públicas" por defecto). Y, por último, siempre es posible cometer errores (por ejemplo, la configuración de la audiencia en las publicaciones de Facebook es rígida, lo que significa que, si una publicación se etiqueta para la audiencia pública, también lo harán todas las siguientes hasta que el usuario vuelva a cambiar la audiencia a la configuración anterior).

Las consecuencias son numerosas: el robo de identidad, la presión social sobre los usuarios (en particular, los adolescentes), la depredación sexual, las consecuencias inesperadas de compartir información (por ejemplo, con un posible empleador), la vigilancia selectiva o, llevado al extremo, la vigilancia masiva de los ciudadanos. Los usuarios no suelen ser conscientes de la precisión de la historia que cuentan sobre sí mismos. En 2009, Le Tigre publicó la vida de Marc L*[97], una historia de varios años de la vida de Marc L*, recogida únicamente de diferentes redes sociales. Más recientemente, la información extraída de las redes sociales ha sido utilizada por la empresa Cambridge Analytica para caracterizar el perfil de los usuarios e influir en su voto mediante mensajes personalizados[98].

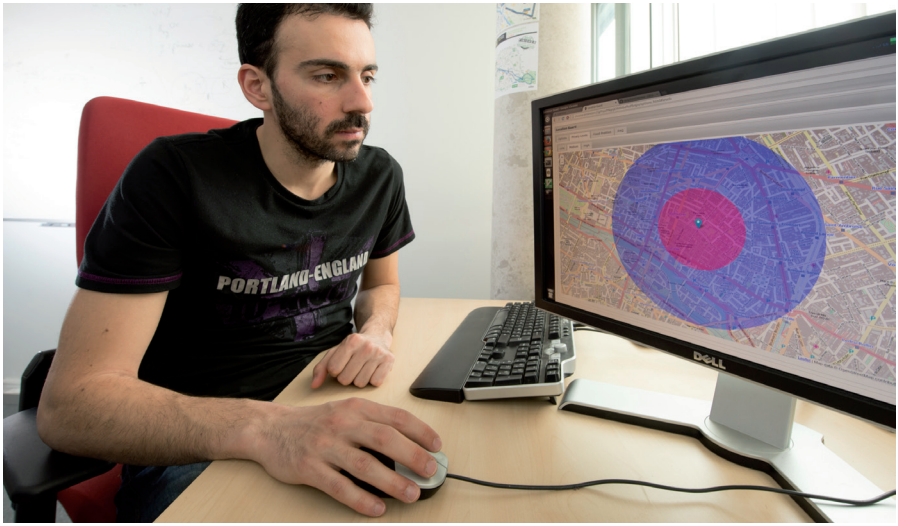
Es necesario investigar para comprender mejor las tendencias, pero también para proteger a los usuarios, en un contexto en el que la minería de datos y el aprendizaje automático se han convertido en herramientas extremadamente potentes.

5.3.2 El lado visible: el caso de la información de geolocalización

La información de geolocalización es otro tipo de información que se comparte con el consentimiento tácito o informado del usuario. Sin embargo, la recogida de la geolocalización de una persona a lo largo del tiempo plantea importantes riesgos: se pueden deducir fácilmente datos personales adicionales (por ejemplo, la ubicación de su casa y su trabajo, o sus hábitos), pero también información personal sensible (por ejemplo, problemas de salud si acude a un hospital o religión si acude regularmente a un lugar de culto). Por tanto, es necesario protegerla, por ejemplo, para beneficiarse de los servicios geolocalizados sin filtrar demasiada información.

Los principales métodos para la protección de la privacidad de los datos de localización se dividen en dos clases: encubrimiento espacial y ofuscación espacial. En el encubrimiento espacial, el objetivo es el anonimato de rastreo para evitar la identificación de un individuo. La mayoría de los métodos de esta clase se basan en el anonimato de grupo, un enfoque muy popular en la literatura sobre el anonimato. La idea general es hacer que los rastros de un individuo sean indistinguibles de los de otros individuos; esto se consigue normalmente informando un área encubierta que es lo suficientemente grande como para contener el tamaño del grupo necesario para cumplir con la restricción de anonimato pretendida. Para limitar el tamaño del área encubierta, algunas propuestas han combinado el encubrimiento

97. http://www.le-tigre.net/Marc-L.html
98. https://ca-political.com/casestudies/casestudydonaldjtrumpforpresident2016



Anonimización de la geolocalización – © Inria / Photo H. Raguet

espacial con el encubrimiento temporal. Cuando el sistema utiliza seudónimos, una necesidad para determinadas aplicaciones, existe el riesgo de que se produzca una vinculación entre puntos pertenecientes a la misma trayectoria del usuario. Para resolver este problema, los investigadores han propuesto las llamadas zonas mixtas, que son zonas en las que se reúnen muchos usuarios y pueden renovar su seudónimo, sin peligro de ser rastreados. Todas las medidas anteriores relacionadas con el encubrimiento espacial necesitan, por supuesto, la intervención de una entidad de confianza que actúe como el servidor de anonimato.

En la segunda clase, la ofuscación espacial, el objetivo es abordar el problema de la identificación de la ubicación del usuario. En general, la privacidad se preserva reduciendo la precisión de la posición. Esto se hace reduciendo la granularidad de la información de localización: el usuario informa una zona en lugar de las coordenadas exactas. Una ventaja importante de este enfoque es que puede realizarse sin la intervención de un tercero de confianza. Sin embargo, este método no es muy robusto, ya que está sujeto a ataques de triangulación: un usuario que envía dos señales consecutivas desde dos zonas distintas revela que está cerca de la frontera entre ellas y con tres señales consecutivas desde zonas distintas revelaría su posición con bastante precisión. Por ello, se han investigado soluciones más eficaces para la ofuscación espacial.

La geo-indistinguibilidad[99][ABCP13] es una de ellas. Extiende la privacidad

99 [ABCP13] Miguel E. Andres, Nicolas E. Bordenabe, Konstantinos Chatzikokolakis y Catuscia Palamidessi. Geo-indistinguibilidad: Differential privacy for location-based systems. En Proceedings of the 2013 ACM SIGSAC Conference on Computer; Communications Security, CCS '13, páginas 901-914, New York, NY, USA, 2013.

diferencial local a métricas arbitrarias con la idea de que la protección de la ubicación del usuario aumenta exponencialmente a medida que disminuye la distancia de la ubicación real. Por lo tanto, un atacante puede determinar que el usuario está en París y no en Londres, y estar razonablemente seguro de que se encuentra en el Barrio Latino, pero no puede decir en qué parte del Barrio Latino exactamente. La geo-indistinguibilidad adopta propiedades interesantes: es independiente del conocimiento colateral del adversario, es robusta con respecto a la composición y no depende de ningún tercero de confianza. Se puede implementar en el extremo del usuario simplemente añadiendo ruido a la ubicación real. Para ello se puede elegir una función de Laplace plana de bajo costo, lo que permite su uso en dispositivos limitados desde el punto de vista computacional, como los smartphones. Gracias a estas propiedades, la geo-indistinguibilidad mediante el mecanismo de Laplace se ha adoptado en varias herramientas para la privacidad de la localización (por ejemplo, LP-Guardian, LP-Doctor y el plugin QGIS de SpatialVision).

5.3.3 El lado visible: el caso de la biometría

La biometría es una potente tecnología para identificar o autentificar a una persona. Los rasgos biométricos, como las huellas dactilares o el iris, son constantes a lo largo del tiempo y altamente diferenciadores; estas son ventajas clave para aplicaciones como la seguridad o el control de acceso. Sin embargo, desde el punto de vista de la privacidad, estas ventajas se convierten en inconvenientes: debido a su estabilidad en el tiempo y a que un individuo no puede cambiar fácilmente sus datos biométricos, la filtración de los rasgos biométricos a una entidad maliciosa da lugar a graves riesgos para la privacidad, incluyendo el seguimiento y el robo de identidad.

Se han propuesto muchas técnicas (como el cifrado, el cifrado homomórfico o el cómputo seguro de múltiples partes) y arquitecturas para tener en cuenta los requisitos de privacidad en la implementación de sistemas biométricos que preservan la privacidad. Algunas soluciones implican elementos criptográficos específicos, como bocetos seguros y bóvedas difusas, otras se basan en adaptaciones de herramientas criptográficas existentes o en el uso de soluciones de hardware seguras. La elección de técnicas concretas y el papel de los componentes (como el servidor central, un módulo seguro, un terminal o una tarjeta inteligente) en la arquitectura tienen un fuerte impacto en las garantías de privacidad ofrecidas.

Teniendo en cuenta la variedad de opciones disponibles y la complejidad de estas técnicas, se han propuesto marcos generales para definir las arquitecturas de privacidad, especificar las distintas opciones, pensar sobre ellas de manera formal y justificar su diseño en términos de hipótesis de confianza.

Los Estados también utilizan cada vez más los sistemas biométricos para proteger los documentos de identidad. Por ejemplo, el gobierno francés autorizó en octubre de 2016 la creación de un archivo centralizado de "documentos

electrónicos seguros" (DES). El principal motivo es la lucha contra la usurpación de la identidad. Sin embargo, el decreto también autorizaba el acceso a la base de datos por parte de una serie de policías y funcionarios. Se han formulado varias críticas sobre los riesgos que un archivo tan centralizado podría suponer para la libertad y la privacidad individuales. El refuerzo de los medios para luchar contra el fraude y la criminalidad, y la exigencia de proteger la privacidad no deberían ser necesariamente incompatibles. Sin embargo, para poder llegar a una decisión sobre los beneficios y los puntos débiles de un sistema de documentos electrónicos seguros, es necesario[100][CM17]:

- definir claramente las funcionalidades deseadas y las ventajas que se pueden esperar de ellas;
- describir las soluciones técnicas de forma suficientemente precisa para permitir su análisis;
- y analizar rigurosamente los riesgos de las violaciones de la privacidad con respecto a los beneficios esperados.

5.3.4 Filtraciones de privacidad ocultas: el caso del rastreo en la web

El despliegue masivo de Internet ha ido acompañado rápidamente de filtraciones de datos personales. Cada visita a un sitio web puede desencadenar una gran variedad de intercambios de datos ocultos a través de múltiples empresas de seguimiento que recogen cada una de ellas grandes cantidades de datos sobre las preferencias y hábitos de los usuarios. La información puede utilizarse entonces para la publicidad dirigida, pero también para discriminar a los usuarios (por ejemplo, mediante precios personalizados) o para la vigilancia.

El seguimiento de la web ha sido posible gracias a la discreta adición de pequeños componentes, llamados rastreadores, a las páginas web. Cada rastreador es propiedad de un tercero, normalmente distinto del propietario del sitio web, y permite a este tercero reconocer a los usuarios en los distintos sitios web que lo incorporan. Estas tecnologías se dividen, a grandes rasgos, en dos tipos: con estado y sin estado. Las técnicas de seguimiento con estado almacenan información en la computadora del usuario que puede ser recuperada posteriormente para reconocerlo. Las cookies de terceros son la técnica de seguimiento en línea con estado más extendida. Para mapear e intercambiar los perfiles de los usuarios, suelen implementarse funciones avanzadas, como la capacidad de reactivar las cookies eliminadas por un usuario o la sincronización de las cookies entre distintas terceras partes.

Por otro lado, las técnicas de rastreo sin estado permiten a terceros reconocer a los usuarios, simplemente por su huella digital, o fingerprinting, sin almacenar nada. La recopilación de varias piezas de información sobre el navegador y el sistema operativo del usuario es suficiente para identificar de forma única cada navegador. En

100 [CM17] C. Castelluccia y D. Le Metayer. Titres electroniques securises : la centralisation des données biométriques est-elle vraiment inévitable ? Nota de análisis Inria, febrero de 2017.

2010, Eckersley demostró por primera vez que la tecnología era muy eficaz a través del proyecto Panopticlick[101]. Trabajos recientes han demostrado que hoy en día la huella digital es tan eficaz en los dispositivos móviles como en las computadoras, en particular gracias a las recientes tecnologías web (por ejemplo, HTML5 trajo atributos altamente diferenciadores) o a los sitios web en los que un usuario está conectado.

Estas prácticas, fuente de importantes filtraciones de privacidad, plantean dos cuestiones complementarias y complejas: ¿cómo detectarlas y cómo protegerse?

En el caso del rastreo web con estado, el bloqueo explícito de las cookies de terceros en la configuración del navegador o la adición de extensiones de bloqueo de anuncios ayudan mucho. Aunque no son perfectas, estas técnicas detectan y desactivan correctamente los rastreadores más comunes e intrusivos[102][MHB+17].

En cuanto a la protección contra el rastreo web sin estado (fingerprinting), desactivar JavaScript es eficiente pero no práctico. La cuestión de la detección automática de rastreadores es compleja y no existe una metodología precisa en la actualidad. La mayoría de las metodologías de detección ignoran muchos rastreadores porque solo comprueban el acceso a las APIs asociadas al fingerprinting en el navegador o aplican un análisis estático muy básico. Las técnicas de seguridad basadas en el lenguaje pueden ir más allá: se pone en marcha un seguimiento para analizar cuánta información es realmente filtrada desde las APIs asociadas al fingerprinting y si esta información es suficiente para identificar a un usuario de forma única. Otro enfoque consiste en añadir diversificación al navegador en los niveles de máquina virtual y API: al introducir suficiente ruido durante el proceso de fingerprinting, los rastreadores pueden ser engañados.

Las anteriores técnicas de protección corresponden a la iniciativa del cliente. Los propietarios de sitios web también pueden estar interesados en proteger a sus usuarios del seguimiento web (por ejemplo, el RGPD hace responsables a los propietarios de sitios web del seguimiento de terceros presente en sus sitios web). Este es el objetivo de las nuevas arquitecturas propuestas, en las que servidores adicionales interceptan y modifican automáticamente las solicitudes web, impidiendo así el seguimiento de terceros.

Por último, hay que tener en cuenta que el modelo de negocio de los sitios web se basa a menudo en la publicidad y que bloquear todo a ciegas perjudica a los proveedores de contenidos. De ahí la cuestión del control del usuario y la elección responsable: en lugar de intentar bloquear todos los rastreadores, autorizar algunos de ellos, por ejemplo los presentes en las páginas web consideradas como menos relevantes.

101. https://panopticlick.eff.org/

102 [MHB+17] Georg Merzdovnik, Markus Huber, Damjan Buhov, Nick Nikiforakis, Sebastian Neuner, Martin Schmiedecker y Edgar R. Weippl. Block me if you can: A large-scale study of tracker-blocking tools. En 2017 IEEE European Symposium on Security and Privacy, EuroS&P 2017,, París, Francia, 26-28 de abril de 2017, páginas 319{333, 2017.

5.3.5 Filtraciones de privacidad ocultas: el mundo inteligente

Más allá del rastreo en la web, la llegada de dispositivos inteligentes y conectados amplió considerablemente las posibilidades de recopilar datos personales, tanto en volumen como en precisión, abarcando ámbitos que hasta ahora estaban fuera de nuestro alcance.

Los teléfonos inteligentes han desempeñado un papel fundamental desde este punto de vista. Estos asistentes personales, fácilmente personalizables con aplicaciones, siempre conectados, equipados con una gran variedad de sensores de alta precisión, contienen y generan mucha información sobre nuestras actividades y núcleos de interés, tanto en el mundo cibernético (Internet) como en el físico. Por ello, los teléfonos inteligentes se han convertido en objetivos ideales. El ecosistema de los teléfonos inteligentes está formado por muchos actores, desde los desarrolladores de aplicaciones hasta los anunciantes. Sin embargo, la recopilación de datos personales la orquestan principalmente las empresas de publicidad y análisis (A&A, Advertising and Analytics)[103], también conocidas como "terceros". En un mundo en el que las aplicaciones gratuitas representan una gran parte de la oferta, las empresas de A&A han desarrollado pequeños softwares de seguimiento que los desarrolladores son invitados a integrar dentro de sus aplicaciones para monetizarlas. Estos rastreadores recogen datos personales de varios miles de millones de sesiones de aplicaciones cada día y los envían a las empresas de A&A para crear perfiles de usuarios cuya precisión aumenta día tras día.

Podría decirse que la recopilación de datos personales a cambio de aplicaciones o servicios gratuitos podría ser aceptable si esta recopilación estuviera documentada en un aviso de política de privacidad, respetando las leyes del país en el que reside el usuario, y si se aplicara de forma "privada por definición" con garantías de minimización de datos y responsabilidad. Pero muchos estudios informan de que ocurre lo contrario. Las empresas de A&A tienden a recoger tanta información personal como sea técnicamente posible[104]. Sin embargo, existen importantes diferencias entre los distintos sistemas operativos (SO), en función de las decisiones tomadas por su editor con respecto a las posibilidades de rastreo del usuario y al control del mismo (por ejemplo, en 2013 Apple fue la primera empresa en prohibir el acceso a los identificadores estables y sustituirlos por un Identificador de Publicidad Específico, bajo pleno control del usuario).

Hoy en día, la moda de los dispositivos corporales tipo "mi yo cuantificado", los electrodomésticos inteligentes, las ciudades inteligentes y los autos conectados, o más generalmente los "dispositivos conectados", permite la recopilación de datos personales en nuevos ámbitos[105]. Esto es incluso más preocupante ya que parte de

103. http://www.mobyaffiliates.com/guides/mobile-advertising-companies/
104. Por ejemplo la empresa InMobi ha sido condenada en 2016 por la FTC (https://www.flc.gov/news-events/press-releases/2016/06/mobile-advertising-network-inmobi-settles-flc-charges-it-tracked) por su uso indebido del permiso ACCESS WIFI STATE Android para rastrear la geolocalización de los usuarios.
105. https://www.cnil.fr/fr/enceintes-intelligentes-des-assistants-vocaux-connectes-votre-vie-privee

los datos que se recogen son información personal sensible y muchos dispositivos conectados siguen siendo muy vulnerables.

Si el modelo de negocio de las empresas que venden dispositivos conectados es diferente al de los desarrolladores de aplicaciones para teléfonos inteligentes, se sabe poco sobre las prácticas reales en términos de confidencialidad e intercambio de datos. Además, el aviso de privacidad que debería acompañar a cada dispositivo conectado para informar al usuario suele faltar o ser ilegible para los no juristas. Por lo tanto, el usuario final es prisionero de un sistema altamente asimétrico.

Es necesario realizar trabajos académicos en estos ámbitos para comprender las distintas facetas del problema, incluidos los modelos de negocio subyacentes. Al dar información transparente de los comportamientos ocultos, al poner de relieve las buenas y malas prácticas, al proponer métodos para mejorar la transparencia y el control de los usuarios, los objetivos finales de estos trabajos son reducir la asimetría informativa del sistema, capacitar a los usuarios y, con suerte, animar a ciertas partes interesadas a cambiar de prácticas.

5.3.6 Fugas de privacidad ocultas: el caso de Internet

La Internet, como conjunto complejo de diversas tecnologías, también presenta riesgos para la privacidad de los usuarios. En esta sección se analizan los riesgos asociados a la red de acceso, al Sistema de Nombres de Dominio (DNS), que es fundamental para casi todas las actividades en Internet, y, por último, a los servicios de protección de la suplantación de identidad que podrían convertirse en herramientas de intrusión de la privacidad.

REDES DE ACCESO INALÁMBRICO

En las redes de acceso inalámbricas, el tráfico que se transmite por enlace inalámbrico suele estar protegido por mecanismos de seguridad como WPA en las redes IEEE 802.11 (wifi). Sin embargo, las cabeceras y el contenido de las estructuras de gestión no están protegidos y, por tanto, están a disposición de los intrusos. La exposición de esta información plantea serias amenazas a la privacidad que se vuelven críticas por la adopción masiva de dispositivos portátiles y el desarrollo de redes inalámbricas.

En concreto, los dispositivos con wifi buscan puntos de acceso cercanos enviando solicitudes de detección. Estas solicitudes pueden incluir el nombre (SSID) de la red a la que el dispositivo ha estado asociado en el pasado. En ese caso, los SSID emitidos por un dispositivo revelan muchos datos personales, como el historial de viajes y la identidad, y también se pueden inferir los vínculos sociales entre los usuarios. Otro aspecto de las tramas 802.11 es la dirección MAC, un identificador único global vinculado al dispositivo. Con este identificador es posible detectar la presencia de personas y seguirlas en el mundo físico.

Las empresas han aprovechado la oportunidad de rastrear a los usuarios de smartphones, lo que ha provocado el rápido crecimiento de los sistemas de seguimiento

Detectando ondas de IoT – © Inria / Photo C. Morel

ciberfísico. Por ejemplo, se despliegan en centros comerciales para medir y analizar los movimientos y hábitos de los clientes, y se ha intentado desplegar estas tecnologías en los equipamientos públicos para mostrar publicidad dirigida. Las medidas adoptadas por estas empresas para supuestamente anonimizar los datos y reducir los riesgos para la privacidad han demostrado tener un impacto limitado.

Por lo tanto, se necesitan soluciones que permitan un análisis que preserve la privacidad para este tipo de aplicaciones. Un enfoque prometedor consiste en utilizar estructuras de datos probabilísticas basadas en filtros de Bloom que incluyan un mecanismo de distorsión para reforzar las garantías de privacidad y, al mismo tiempo, permitir la estimación precisa del número de identificadores de dispositivos detectados. En respuesta a los problemas de rastreo, los proveedores de wifi empezaron a implementar la aleatorización de la dirección MAC, una técnica en la que un seudónimo aleatorio y temporal sustituye a la dirección MAC real en las tramas IEEE 802.11. A pesar de la adopción de esta tecnología de mejora de la privacidad, los estudios han demostrado que todavía es posible rastrear a los usuarios mediante ataques activos que obligan a los dispositivos a revelar su dirección MAC real y mediante ataques de huella digital basados en el contenido y en los tiempos de las solicitudes de detección que podrían utilizarse para identificar dispositivos.

Queda mucho por hacer en el caso de la wifi y otras tecnologías inalámbricas

similares para reducir los riesgos para la privacidad asociados a su uso; este sigue siendo un tema de investigación activo.

SERVICIOS BÁSICOS DE INTERNET

Más allá de las redes de acceso, el protocolo DNS, fundamental en Internet (véase 2.2), sí presenta amenazas a la privacidad que siguen sin resolverse. La transición progresiva a la versión "DNS Security Extensions" (DNSSEC) resuelve algunas de las amenazas de seguridad relacionadas, pero no aborda los requisitos de confidencialidad. Por ejemplo, incluso en presencia de DNSSEC, un intruso presente entre un cliente y su gestor de DNS podrá analizar el tráfico del DNS e identificar la mayoría de los sitios web visitados, incluso en caso de tráfico HTTPS cifrado en la Web[106]. Esta información filtrada es considerada como datos personales (está asociada a una persona física) y da mucha información sobre los puntos de interés del usuario. Otros riesgos se discuten en[107][Bor15].

Se están estudiando varias líneas en el contexto del grupo de trabajo[108] del IETF "DNS PRIVate Exchange" (DPRIVE), aprovechando las comunicaciones cifradas (por ejemplo, a través de TLS o DTLS). La investigación sigue siendo necesaria, ya que las consideraciones prácticas hacen que la situación en el mundo real sea más compleja de lo que parece.

SISTEMAS DE DETECCIÓN DE ACTIVIDADES MALICIOSAS

La navegación por la web puede llevar a los usuarios a visitar sitios web maliciosos. Las técnicas de navegación segura se han creado para detectar las URL contenidas en una lista negra (por ejemplo, sitios web que se sabe que están implicados en ataques de phishing o que contienen malware). Aunque son muy útiles para el usuario final, estos servicios también son potencialmente peligrosos desde el punto de vista de la privacidad. Almacenar toda la lista de URLs de una lista negra en la computadora es imposible, y confiar en un servicio externo para comprobar las URLs visitadas es demasiado intrusivo (el proveedor del servicio estaría en posición de recoger todas las URLs visitadas por un usuario). Por tanto, en la práctica se utilizan soluciones intermedias, y es necesario analizarlas cuidadosamente desde el punto de vista de la privacidad.

El servicio de Búsqueda Segura de Google o Safe Browsing, a su vez reutilizado por la mayoría de los navegadores web, ha sido estudiado desde la perspectiva de la privacidad. Aunque la empresa probablemente hizo todo lo posible para anonimizar los datos y cumplir con la privacidad, este servicio de Google carece de transparencia y responsabilidad. En general, sería muy beneficioso exponer y debatir su tecnología con terceros independientes y de confianza.

106. 4El tráfico de 4DNS entre los clientes y sus servidores DNS pasa tradicionalmente por comunicaciones UDP, incompatibles con la protección TLS.
107. [Bor15] Stephane Bortzmeyer. DNS privacy considerations. *RFC*, 7626, 2015. https://tools.ietf.org/html/rfc7626
108. https://datatracker.ietf.org/wg/dprive/about/

[Desafío de investigación 9] Hacia un mundo conectado inteligente que preserve la privacidad

Nuestro mundo conectado experimenta un crecimiento sin precedentes en cuanto a la recogida de datos personales, con prácticas cada vez más intrusivas para la intimidad del ciudadano. Navegar por la web, utilizar smartphones y otros dispositivos inteligentes, conducir un coche conectado -y pronto autónomo- son actividades que generan filtraciones de datos personales. La falta de transparencia (muchos servicios y dispositivos se comportan como cajas negras) y la falta de control del usuario (cómo expresar su consentimiento u oposición cuando no hay información, ni interfaz de usuario) son problemas importantes. El reconocimiento de estos comportamientos ocultos se ve dificultada por el número y la complejidad de las tecnologías subyacentes específicas de cada entorno. Por ejemplo, la identificación de las prácticas de rastreo en una página web requiere análisis avanzados de ejecución de JavaScript, mientras que la supervisión de las aplicaciones de los teléfonos inteligentes necesita sistemas específicos, y la supervisión de determinadas tecnologías de comunicación inalámbrica sigue sin resolverse en su mayor parte. El análisis de estos flujos de datos es necesario para evaluar posibles fugas de privacidad, por ejemplo, en un hogar inteligente. Estas desafiantes y diversas actividades de investigación son esenciales para aportar transparencia, poner de relieve las buenas y malas prácticas y permitir a los reguladores hacer cumplir las leyes de protección de datos. De este modo, esta investigación contribuye directamente a dar forma a nuestro futuro mundo conectado e inteligente.

[Equipos Inria] Análisis de la privacidad de los sistemas existentes

↗ El equipo **CIDRE** trabaja en el control de los usuarios en el contexto de las redes sociales. El equipo ha propuesto una clasificación de las redes sociales existentes basada en el tipo de implementación (centralizada o distribuida) de sus funcionalidades (por ejemplo, comunicación, búsqueda o almacenamiento).

El equipo también trabajó en la privacidad de la localización, produciendo la herramienta GEPETO[a], cuyo objetivo es permitir al usuario diseñar, ajustar, experimentar y evaluar varios algoritmos de limpieza y ataques de inferencia y evaluar los intercambios resultantes entre privacidad y utilidad.

↗ El equipo **COMETE** trabaja en la privacidad de la localización, proponiendo la noción de geo-indistinguibilidad, [ABCP13], que extiende la privacidad diferencial a las métricas de distancia: la localización precisa de un dispositivo se protege mediante la adición de ruido controlado a la posición notificada. El equipo también ha desarrollado una herramienta basada en la geo-indistinguibilidad, llamada Location Guard[b], Guardián de la Localización, una extensión del navegador que permite proteger la ubicación del usuario mientras accede a sitios web con reconocimiento de la ubicación, añadiendo ruido controlado a la misma.

↗ El equipo **DIVERSE** trabaja en el ámbito de la huella digital de los navegadores web. El sitio web Am I unique[c] demuestra cómo ha evolucionado la situación con las tecnologías

web más recientes (por ejemplo, HTML5 aportó atributos altamente diferenciadores) y muestra que el reconocimiento de la huella digital es hoy en día tan efectivo en los dispositivos móviles como en las computadoras, aunque por razones distintas. Para protegerse del fingerprinting, el equipo también trabaja en la diversificación de los navegadores, con la prueba de concepto FPRandom[d].

↗ El equipo **INDES** trabaja en el análisis de diversas formas de rastreo en aplicaciones web. En un trabajo conjunto con **PRIVATICS**, el equipo también demostró que los navegadores web pueden ser objeto de reconocimiento de huella digital a través de las extensiones[e] que el usuario instala y los sitios web donde se registran. Por último, el equipo propone una técnica de complemento del servidor para protegerse contra el rastreo web: dos servidores adicionales reescriben y redirigen respectivamente las peticiones originales de la aplicación web para evitar el rastreo de terceros automáticamente[f].

↗ El equipo **PESTO** trabaja en la privacidad en las redes sociales en línea, con el objetivo de informar a los usuarios de la información latente que puede inferirse de su información publicada.

↗ El equipo **PRIVATICS** trabaja en las amenazas a la privacidad introducidas por la sociedad de la información, en particular para la biometría, el rastreo web, los teléfonos inteligentes, el mundo inteligente (IoT), las redes de acceso inalámbricas o los sistemas de detección de sitios web maliciosos. Los objetivos son comprender la situación, analizar las amenazas y, si procede, diseñar soluciones que preserven la privacidad para prevenirlas o mitigarlas. **Por ejemplo, el equipo propuso un marco general para la concreción y el razonamiento formal de arquitecturas biométricas, y lo aplicó a varias arquitecturas para el control de acceso biométrico. El equipo también contribuyó en [CM17] al debate sobre los sistemas de documentos electrónicos seguros (TES). A través del proyecto MyTrackingChoices, el equipo propuso un nuevo tipo de bloqueador de anuncios que permite al usuario elegir qué categorías de sitios web pueden rastrearlo o no. El equipo también trabajó en el tema de la privacidad de los teléfonos y el mundo inteligente (IoT), siguiendo un enfoque multidisciplinar con juristas y economistas, y en las cuestiones de transparencia y control del usuario dentro de una ciudad inteligente. Por último, el equipo fue pionero en el problema de las filtraciones de privacidad dentro de las redes de acceso inalámbrico y desarrolló las herramientas de escaneo/rastreo de wifi Wifiscanner[g] y Wombat[h].**

↗ El equipo **SPIRALS** trabaja en el análisis de las huellas de los navegadores web, en particular su incongruencia y evolución en el tiempo.

a. https://gforge.inria.fr/projects/gepeto/
b. https://github.com/chatziko/location-guard
c. https://amiunique.org/
d. https://github.com/plaperdr/fprandom
e. https://extensions.inrialpes.fr/
f. https://www-sop.inria.fr/members/Doliere.Some/essos/
g. https://github.com/cunchem/gtk-wifiscanner
h. https://github.com/Perdu/wombat

6

Infraestructuras críticas, sistemas y aplicaciones: casos reales para la seguridad

En este capítulo examinamos la ciberseguridad bajo el ángulo de las aplicaciones. Nos centramos en algunos entornos seleccionados en los que es probable que la ciberseguridad tenga un fuerte impacto. Seleccionamos en particular dominios sensibles a la seguridad:

- infraestructuras: la nube, las redes definidas por software (SDN) y la cadena de bloques (blockchain);
- sistemas críticos y ciberfísicos: la Internet de las cosas (IoT) y los sistemas industriales;
- áreas de aplicación: medicina, robótica (incluidos los vehículos autónomos conectados) y aprendizaje automático (Machine Learning, o ML).

Esta lista no es, por supuesto, exhaustiva, pero indica muy bien que muchos sistemas y áreas de aplicación son críticos para la seguridad y que las consideraciones de seguridad deben incluirse necesariamente en el proceso de diseño. La clasificación en infraestructuras críticas, sistemas críticos y áreas de aplicaciones críticas tampoco es absoluta –por ejemplo, los implantes médicos son sistemas ciberfísicos– y las aplicaciones críticas pueden, por supuesto, basarse en infraestructuras o sistemas críticos.

6.1 Infraestructuras críticas

Las infraestructuras de comunicación están diseñadas principalmente para ofrecer un servicio, a menudo con la facilidad de uso y la eficiencia como objetivos principales. Sin embargo, la seguridad también es crucial, ya que estas infraestructuras pueden utilizarse para almacenar y manipular datos sensibles. La pérdida de disponibilidad (o simplemente de eficiencia) y de integridad de los datos también puede tener un fuerte impacto económico. Nos centraremos en tres ejemplos de este tipo de infraestructuras críticas: la nube, las redes definidas por software y la cadena de bloques.

6.1.1 Seguridad y privacidad en la nube

[Resumen]
La seguridad sigue siendo un obstáculo para conseguir una mayor adopción de los servicios en la nube por parte de las organizaciones para sus servicios críticos. La privacidad también se ha convertido en una cuestión importante para los usuarios de la nube en la era de la IoT. Por lo tanto, un reto importante para los proveedores de la nube es proporcionar a sus usuarios tanto servicios de seguridad con garantías asociadas en los acuerdos de nivel de servicio (SLA) como servicios de computación y almacenamiento de datos que preserven la privacidad. Los problemas de seguridad y privacidad se agravan en el contexto de las nubes descentralizadas, es decir, la computación de proximidad y la computación presente entre el usuario y la nube (fog computing). La confianza y otros supuestos de seguridad varían con el tiempo debido a los cambios en la tecnología y al descubrimiento de nuevas vulnerabilidades. Estas evoluciones han

conducido necesariamente a una evolución adecuada de los modelos de amenaza en la seguridad de la nube, considerando no sólo los ataques al sistema operativo del cliente, sino también los ataques al monitor de máquinas virtuales (hypervisor). La seguridad de los hipervisores y de los sistemas operativos y la acreditación de las propiedades de seguridad en ellos parecen ser temas de investigación oportunos e importantes.

Vehículos autónomos – © Inria / Photo G. Scagnelli

La seguridad es una de las principales preocupaciones en la adopción del modelo en la nube[109]. La infraestructura de hardware y de software de bajo nivel es propiedad del proveedor de la nube y está gobernada por él, mientras que los clientes que externalizan su sistema de información tienen el control sobre los servicios y el sistema operativo que se ejecuta en sus infraestructuras virtualizadas. La virtualización de servidores permite la ejecución de diferentes sistemas operativos y/o aplicaciones de diferentes clientes en el mismo servidor de un centro de datos. Esta situación de multi-arrendamiento conlleva amenazas específicas de seguridad y privacidad.

Los entornos en la nube se enfrentan a múltiples amenazas de seguridad que se originan en diferentes niveles de permisos (niveles de aplicación, red y sistema operativo) en la infraestructura. En un entorno de nube IaaS (Infraestructura como Servicio), la superficie de ataque se amplía con la adición del hipervisor de alto nivel de permiso de acceso, como el bloque de construcción de una infraestructura

109. http://www.infosecbuddy.com/download-cloud-security-report

en la nube, así como su API de gestión expuesta a la web. En este contexto, la seguridad afecta a dos actores diferentes: los clientes y los proveedores.

A los clientes les preocupa la seguridad de sus activos externalizados, especialmente si están expuestos a Internet. Los ataques dirigidos a los sistemas de información tradicionales también podrían dirigirse a las aplicaciones que se ejecutan dentro de las máquinas virtuales en una infraestructura subcontratada.

El proveedor también está preocupado por la seguridad de la infraestructura subyacente, especialmente porque no tiene conocimiento de las aplicaciones alojadas en ella y su carga de trabajo. En un entorno de nube, las amenazas a la seguridad procedentes de clientes dañinos contra otros clientes legítimos y sus recursos, las amenazas contra la infraestructura del proveedor, así como las amenazas hacia la API del proveedor, deben ser consideradas. Los proveedores de la nube deben evaluar los riesgos de seguridad teniendo en cuenta las infraestructuras virtualizadas alojadas y las vulnerabilidades de las tecnologías de virtualización. Necesitan saber qué máquinas virtuales pueden interactuar a través de protocolos de red.

En las nubes IaaS, el proveedor de la nube gestiona la infraestructura de supervisión de la seguridad de la nube, mientras que los inquilinos de la nube gestionan su sistema de información subcontratado. La supervisión de la seguridad es esencial en las nubes. Un reto específico es garantizar que la infraestructura de supervisión de la seguridad se reconfigura automáticamente cuando se producen eventos dinámicos en la nube (por ejemplo, la migración de máquinas virtuales). Se alienta a los clientes a confiar en las afirmaciones del proveedor (por ejemplo, la disponibilidad de la infraestructura) gracias a la garantía dada por los Acuerdos de Nivel de Servicio (SLA), haciendo un intercambio entre la información privada revelada por el inquilino y el servicio de monitoreo ofrecido.

Aunque el multi-arrendamiento maximiza la eficiencia de los recursos del proveedor de la nube, también se ofrece la posibilidad de que la máquina virtual de un arrendatario pueda estar ubicada en la misma máquina física que una máquina virtual maliciosa. Esto a su vez engendra una nueva amenaza: quebrantar la protección de recursos proporcionados por el hipervisor y el hardware y obtener acceso a datos no autorizados o perturbar el funcionamiento de las máquinas virtuales legítimas. Uno de los ataques más destacados que refleja esta amenaza es el ataque de canal lateral, en el que un adversario con una máquina virtual asignada accede a información perteneciente a otras máquinas virtuales (por ejemplo, contraseñas o claves criptográficas). Por ejemplo, los atacantes podrían utilizar las cachés de CPU compartidas como canales laterales para extraer información sensible de una máquina virtual asignada.

Normalmente se supone que los proveedores de la nube son "honestos pero curiosos": dada su posición privilegiada con respecto a sus clientes, un proveedor malicioso o un proveedor cuyo sistema se vea comprometido podría amenazar

la privacidad de sus arrendatarios. Los mecanismos de vigilancia de máquinas virtuales pueden ser utilizados por el proveedor de la nube para supervisar la infraestructura virtual de los clientes. Externalizar los datos en la nube significa delegar el control de los datos en el proveedor. Los clientes de la nube tienen poco o ningún control sobre dónde y durante cuánto tiempo se almacenan los datos y a qué terceros se remiten. Además, el proveedor de la nube puede poner a disposición recursos ubicados en diferentes países con distintas regulaciones de los datos, lo que resulta en una protección que depende del lugar donde se almacena la información, y esto es a menudo transparente para los clientes de la nube.

[Equipos Inria] Seguridad y privacidad en la nube

↗ El equipo **AVALON** está interesado en el diseño de las propiedades de seguridad de las aplicaciones y su implementación automática para las aplicaciones ejecutadas en las nubes. Propusieron un enfoque orientado a la especificación en el que la seguridad se expresa como propiedades en un lenguaje ajeno al sistema para facilitar la expresión de los requisitos del usuario tanto para la aplicación como para su seguridad, lo que resulta en el conjunto de herramientas Sam4C. También desarrollaron mecanismos para proporcionar el despliegue automático de la aplicación y la ejecución automática de su seguridad, con propiedades contrastadas.

↗ El equipo **CASCADE** trabaja en la privacidad en la nube, diseñando una nueva generación de protocolos de computación segura de múltiples partes que permiten el procesamiento de datos cifrados.

↗ El equipo **CIDRE** investiga la evaluación de la seguridad en las nubes teniendo en cuenta las tecnologías de virtualización, hipervisores y SDN, en la extracción de la conectividad en las infraestructuras de la nube y las vulnerabilidades específicas de la nube mediante la generación de grafos de ataque.

↗ El equipo **MYRIADS** tiene como objetivo integrar los términos de supervisión de la seguridad en los acuerdos de nivel de servicio de las nubes IaaS y diseñar una infraestructura de supervisión de la seguridad en la nube autoadaptable. El equipo ha diseñado un entorno para esta finalidad que es capaz de alterar la configuración de sus componentes y adaptar la cantidad de recursos computacionales disponibles en función del tipo de evento dinámico que se produzca en una infraestructura de nube.

↗ El equipo **STACK** ha propuesto un procedimiento integrador para la formulación informativa y correcta de aplicaciones que preservan la privacidad en la nube. El enfoque propuesto proporciona el soporte de lenguaje para la integración de tres técnicas: cifrado simétrico, fragmentación vertical de datos y procesos de lado del cliente para hacer que las aplicaciones en la Nube preserven la privacidad. El equipo también estudia las amenazas de aislamiento en entornos de contenedores investigando los ataques de canal lateral en el contexto de las nubes de borde.

6.1.2 Seguridad de las redes definidas por software (SDN)

[Resumen]
La softwarización de las redes es la evolución actual en este campo. El objetivo es mejorar la flexibilidad y reducir los costos. Sin embargo, estas evoluciones centralizan el control de una red, proporcionando un único punto débil para el atacante. Además, la softwarización de la red permite un mejor acoplamiento entre la red y las aplicaciones gracias a diversas API. Aunque antes estaban restringidas a un número limitado de actores, estas API facilitan significativamente el desarrollo de aplicaciones, pero también crean una nueva superficie de ataque.

Las tecnologías de red han evolucionado mucho en los últimos tiempos. El cambio hacia la softwarización de las redes ha modificado profundamente las arquitecturas y operaciones de red. Los paradigmas emergentes más notables son SDN (Software-Defined Networking, o redes definidas por software y NFV (Network Function Virtualization, o virtualización de funciones de red). La SDN propugna un controlador lógicamente centralizado y potente que sustituya a los algoritmos dispersos (la antigua panacea), dejando únicamente los fines de transmisión a los dispositivos de red. El NFV permite ejecutar cualquier tipo de función de red como una máquina virtual, asumiendo algún controlador central o nivel de gestión para organizar el despliegue de la función. Por tanto, está bien alineada con la cloudificación de todo. Aunque las redes centrales seguirán dependiendo de hardware de red específico de gama alta, la softwarización de las redes se basa naturalmente en servidores estándar en una arquitectura similar a la de los centros de datos. El objetivo final es mejorar la flexibilidad de las operaciones de red y reducir los costos de forma sustancial. Sin embargo, este cambio de paradigma conlleva retos inherentes a la seguridad.

La centralización del control de una red reduce el enfoque del atacante y aumenta la vulnerabilidad: El atacante puede interrumpir eficientemente una red comprometiendo un controlador central, mientras que los algoritmos dispersos son más robustos contra los ataques individuales. La denegación de servicio dirigida contra un solo controlador puede provocar un grave impacto. Las nuevas tecnologías, los nuevos marcos y los nuevos protocolos vienen acompañados de estos nuevos paradigmas para integrar el mayor número posible de funcionalidades, aumentando así la superficie de ataque. Garantizar la seguridad y detectar el uso indebido de estos protocolos es de suma importancia.

Además, la softwarización de las redes permitirá un mejor acoplamiento entre la red y las aplicaciones gracias a diversas API. El desarrollo de aplicaciones ya no estará restringido a un número muy limitado de actores, como ocurre actualmente con los dispositivos de hardware producidos por muy pocos fabricantes. El estricto control de calidad deberá responder a las exigencias de usabilidad y flexibilidad del mercado. Por lo tanto, el comportamiento de las aplicaciones en fase de ejecución puede conducir,

por error o de forma intencionada, a violaciones abiertas de la seguridad en la red. Además, estas brechas pueden ser consecuencia de un conjunto de aplicaciones que solo son problemáticas cuando se utilizan conjuntamente. La verificación de las políticas y la detección de las anomalías de configuración en las redes softwarizadas deben tener en cuenta que las configuraciones son dinámicas, por lo que utilizar sólo el análisis estático no es apropiado. El encadenamiento de funciones de servicio (SFC, o Service Function Chaining) consiste en encadenar múltiples funciones de red virtualizadas –FRV– (denominadas VNF o virtualized network functions en inglés) para ofrecer servicios innovadores, incluyendo, por ejemplo, servicios de seguridad (firewall, detección de intrusiones, inspección profunda de paquetes, proxy, etc.); varias cadenas podrían incluso compartir algunas FRV.

La FRV permite una mayor flexibilidad y debería hacer que las funciones de red sean más elásticas. Sin embargo, las funciones de red pueden colocarse en las propias máquinas físicas. Dado que la idea central es utilizar las funciones de red del mercado, es difícil conocer o predecir su comportamiento exacto. Además, los cortos ciclos de desarrollo y despliegue previstos pueden ser la fuente de comportamientos dañinos, introducidos por error o de forma intencionada. Comprobar y evaluar la respuesta de una función de red desde el punto de vista de la seguridad es esencial. Se pueden aprovechar varios enfoques: análisis de código estático si es posible, análisis dinámico o evaluaciones. Además, el aislamiento entre las funciones de red que se encuentran alojadas en un mismo lugar es primordial. Mientras que el aislamiento de las máquinas virtuales se ha estudiado ampliamente en el pasado para hacer sostenible el modelo de la nube, el juego cambia cuando se trata de virtualizar las funciones de red. Además, la consecución de operaciones a la velocidad de transmisión de datos en línea conduce a la relajación de las propiedades de aislamiento sólidas y, por tanto, puede dar lugar a problemas de seguridad. Es necesario encontrar un buen equilibrio entre seguridad y rendimiento, que podría ajustarse en función de la relevancia de los FRV. Es necesario detectar y evitar que una FRV que se comporte mal afecte al funcionamiento de una FRV alojada. Del mismo modo, el despliegue de FRV puede integrar restricciones de seguridad en lugar de ser visto como un puro problema de asignación de recursos.

[Equipos Inria] Seguridad de las redes definidas por el Software

↗ El equipo **COATI** se centra en la optimización de los recursos utilizados por los dispositivos de reenvío en la SDN. El objetivo es poder almacenar más reglas de transferencia de forma comprimida manteniendo un rendimiento aceptable y una carga de control limitada.

↗ El equipo **DIANA** contribuye a la definición de nuevas soluciones para los paradigmas SDN y FRV que tienen en cuenta las consideraciones de seguridad. Además, el equipo evalúa el rendimiento de las tecnologías relacionadas con ellos y, por lo tanto,

indirectamente, su capacidad para resistir las fuertes presiones creadas por los ataques de denegación de servicio.
↗ El equipo **RESIST** trabaja en la elaboración y las pruebas de nuevas definiciones de programación de redes. El tiempo de funcionamiento y la ejecución en el servidor local de los FRV necesitan una nueva representación del mapa de datos. Dado que el aislamiento entre FRV puede no estar garantizado, se hace necesario introducir mecanismos de protección y métodos para predecir el tiempo de ejecución y las interdependencias.
↗ El equipo **RESIST** y **VERIDIS** tienen como objetivo verificar los SFC transformando políticas complejas en una representación formal. La verificación es necesaria para detectar posibles incoherencias debidas al despliegue descentralizado de múltiples SFC.

6.1.3 Cadenas de bloques (Blockchain)

[Resumen]
Al presentar un libro de registros fiable, sólo adjuntable como anexo e inmutable, con varios modelos para escribirlo y gestionarlo (totalmente descentralizado, desregulado, regulado, en consorcio, etc.), las cadenas de bloques permiten muchas aplicaciones que dependen de esta nueva característica de seguridad y sus infraestructuras relacionadas. Sin embargo, su seguridad real y su nivel de confianza deben ser debidamente corroborados con análisis tanto de las comunidades de criptografía como de sistemas dispersos. Además, como ocurre con cualquier sistema de Internet, la s cadenas de bloques pueden sufrir, por ejemplo, ataques de red de bajo nivel, errores de software o fallos, que pueden aparecer cuando las cadenas de bloques se utilizan para aplicaciones de alto nivel (contratos inteligentes). Otras características adicionales, como la fuerte privacidad y el anonimato, también pueden entrar en conflicto con los requisitos de seguridad de los organismos públicos y legales.

UNA PROMESA DE SEGURIDAD

La tecnología blockchain, bastante de moda y emergente, puesta en marcha por primera vez por el protocolo Bitcoin, es una promesa de seguridad. Una cadena de bloques, en cualquiera de sus variantes, implementa un libro de registros electrónico seguro, una funcionalidad similar a los libros de registros existentes gestionados por bancos, notarios, estados, etc., proporcionando la principal propiedad de seguridad de que los datos registrados en el libro no pueden ser eliminados o modificados. Efectivamente, la cadena de bloques proporciona la integridad del historial, basándose en la noción de funciones hash criptográficas: el hash del último bloque confiable certifica la integridad de todo el libro de registros desde su inicio.

SEGURIDAD DEL LIBRO DE REGISTROS

Este libro de registros puede ser garantizado por una única entidad centralizada, en la que se confía para certificar el último bloque de datos de forma sistemática[110][HS91] y para publicar regularmente las sumas de comprobación de integridad concatenadas del libro de contabilidad. Descentralizar la tarea de certificar la cadena de bloques es la principal innovación de Bitcoin. La descentralización podría seguir un modelo entre pares, con nodos no registrados, desconocidos y no fiables (es decir, sin permisos asignados como en Bitcoin o Ethereum). La descentralización también podría seguir un modelo con permisos asignados, en el que o bien se sabe que algunos participantes tienen la función de firmar el último bloque mientras se controlan entre sí (por ejemplo, Hyperledger, blockchains industriales), o bien los participantes son agentes involucrados anónimos que se considera que certifican el libro de registros de forma honesta para proteger su participación.

En consecuencia, el análisis de la seguridad de una cadena de bloques depende del modelo social y político subyacente a la cadena de bloques: la cadena de bloques de Bitcoin puede ser atacada por un enemigo poderoso que controle el 51% de la potencia de cálculo de la red de Bitcoin (Namecoin ha mostrado esta debilidad[111][ANSF16]); las cadenas de bloques de consorcios pueden reescribirse completa y rápidamente si las claves de firma se ven comprometidas; y las cadenas de bloques con participación se enfrentan al riesgo de que una parte involucrada no racional actúe en contra de sus propios intereses financieros.

DINÁMICA DEL PROTOCOLO

Además de las propiedades de integridad del libro de registros, el problema de determinar dinámicamente el bloque de datos concreto que debe añadirse al libro de registros pone de manifiesto el dominio de los algoritmos dispersos, donde las cuestiones de actualidad y corrección son fundamentales. Dado que la cadena de bloques se replica entre los participantes, es esencial que todos tengan la misma visión de ella: si no lo hacen, un atacante puede aprovecharse de las divergencias en la cadena de bloques. Aunque la mayoría de los protocolos de blockchain establecen que las posiciones son las mismas después de un cierto periodo de tiempo, siempre hay un periodo de inestabilidad antes de que se alcance el acuerdo. Por lo tanto, pueden surgir ataques (por ejemplo, el doble gasto) y es necesario abordar los problemas no criptográficos de los algoritmos descentralizados.

110. [HS91] Stuart Haber y W. Scott Stornetta. How to time-stamp a digital document. Journal of Cryptology, 3(2):99-111, enero de 1991

111. [ANSF16] Muneeb Ali, Jude Nelson, Ryan Shea y Michael J. Freedman. Blockstack: A global naming and storage system secured by blockchains. En 2016 USENIX Annual Technical Conference (USENIX ATC 16), páginas 181-194, Denver, CO, 2016. Asociación USENIX.

ATAQUES A LA RED

Los problemas de red no suelen tenerse en cuenta, pero las cadenas de bloques pueden sufrir ataques de red de bajo nivel. Su seguridad no puede ser garantizada únicamente por el propio protocolo criptográfico. Recientemente, se ha demostrado que la potencia de minería de las redes de bitcoin está muy desequilibrada y que es posible un ataque a nivel de red a través de la partición de la red superpuesta (secuestro del Border Gateway Protocol , o BGP)[112][AZV16]. Todavía no se ha realizado una evaluación exhaustiva de los ataques a nivel de red y su impacto en la seguridad y el rendimiento de la cadena de bloques. Es necesario establecer un modelo de evaluación de amenazas y adaptarlo a cada tecnología de blockchain. La seguridad puede mejorarse añadiendo mecanismos de seguridad en el diseño (requisitos a nivel de red, como restricciones relativas a la topología) o contramedidas durante el tiempo de ejecución (acoplamiento entre la red y la cadena de bloques). En primer lugar, se necesita un esquema de supervisión para evaluar la seguridad de una cadena de bloques. En segundo lugar, pueden aplicarse dinámicamente mecanismos como la creación de listas negras de nodos o el refuerzo del modelo de consenso (serán útiles las redes programables, con configuraciones flexibles). Las políticas de red específicas de la cadena de bloques podrían definirse, desplegarse y verificarse automáticamente, por ejemplo, asumiendo la cooperación entre sistemas autónomos que intercambian información de rutas BGP.

FALLOS DE SOFTWARE

Un sistema de cadena de bloques tiene que ser implementado sobre software, al igual que las aplicaciones que utilizan cadenas de bloques. Estas implementaciones pueden encontrar los mismos problemas que cualquier sistema, como errores de programación, ataques, comportamientos no deseados, etc. Sin embargo, estos problemas son mucho más críticos para las cadenas de bloques, porque sus aplicaciones manejan divisas y valores. Esto significa que cualquier problema tendrá un impacto financiero directo, dando a los cibercriminales un fuerte incentivo económico para encontrar o crear tales problemas. Además, las aplicaciones deben hacer frente a características notoriamente propensas a errores como la ejecución simultánea, la distribución, la autoridad y el secreto. Aunque no se han producido problemas importantes con los sistemas Bitcoin o Ethereum en sí, tanto los clientes (la bolsa MtGox) como las aplicaciones (el DAO) han sufrido vulneraciones devastadoras.

BLOCKCHAINS COMO LADRILLOS DE CONSTRUCCIÓN PARA PROTOCOLOS DE NIVEL SUPERIOR

Una cadena de bloques puede utilizarse como un ladrillo de construcción que

112. [AZV16] Maria Apostolaki, Aviv Zohar y Laurent Vanbever. Hijacking bitcoin: Large-scale network attacks on cryptocurrencies. CoRR, abs/1605.07524, 2016.

proporciona una base, donde se pueden implementar programas y protocolos de nivel superior. Por ejemplo, Ethereum concibe la cadena de bloques como un almacenamiento con integridad para los datos y los programas, donde los programas pueden activarse para escribir y actualizar los datos en la cadena de bloques, sin borrar las versiones anteriores. Del mismo modo, Bitcoin tiene una noción débil, pero útil, de dinero programable que permite, por ejemplo, las transacciones fuera de la cadena y los intercambios atómicos entre cadenas. Incluso si se considera la cadena de bloques subyacente como un libro de registros criptográfico perfecto, el desarrollo de programas y protocolos de alto nivel tiene los mismos problemas que el desarrollo de protocolos criptográficos de alto nivel a partir de bloques básicos simples. Debido a que la privacidad es inexistente por defecto en la cadena de bloques sin permisos, se han hecho algunos esfuerzos para garantizar la privacidad utilizando pruebas de nulo conocimiento y otra criptografía avanzada (Monero, Zcash). En este punto, los requisitos de control o las restricciones legales, por ejemplo, KYC (Know Your Customer), chocan con la privacidad proporcionada por la criptografía fuerte, pero este conflicto a veces puede ser solventado, dependiendo del modelo de blockchain.

[Equipos Inria] Blockchain

↗ El equipo **ANTIQUE** investiga sobre la semántica, la representación, la prueba formal y el análisis automático de sistemas, junto con la verificación de las propiedades de seguridad, en estructuras de datos descentralizadas y protocolos de consenso.

↗ El equipo **AVIZ** desarrolla herramientas de visualización y métodos de exploración para el análisis interactivo de conjuntos de datos complejos y de gran tamaño, en particular el blockchain de Bitcoin.

↗ El equipo **CASCADE** se dedica a la criptografía con rigurosas pruebas de seguridad y es experto en protocolos para dinero electrónico (“criptodivisas” centralizadas y descentralizadas) y herramientas de privacidad para blockchains.

↗ El equipo **CIDRE** está diseñando algoritmos de acuerdo distribuidos para mejorar la seguridad, el rendimiento y la escalabilidad de las cadenas de bloques sin permisos, centrándose tanto en las redes subyacentes entre pares como en la estructuración de la información en las cadenas de bloques.

↗ El equipo **COAST** está investigando sobre el intercambio seguro de datos entre pares y la elaboración de servicios para entornos de colaboración sin una autoridad central.

↗ El equipo **DELYS** investiga sobre sistemas dispersos, desde la teoría hasta los algoritmos y las implementaciones, incluyendo el blockchain como mecanismo central en el que trabajan para mejorar su rendimiento y escalabilidad mediante la reordenación de las operaciones sin comprometer la seguridad y la interoperación entre blockchains.

↗ El equipo **GRACE** estudia la privacidad y los protocolos seguros con múltiples partes y cómo pueden utilizarse en el contexto de las cadenas de bloques, así como la teoría de la codificación para el almacenamiento descentralizado.

↗ El equipo **PROSECCO** trabaja en el análisis, diseño e implementación de protocolos de seguridad basados en la tecnología blockchain, por ejemplo, contratos inteligentes, mediante el uso de métodos formales, lenguajes y herramientas de software.
↗ El equipo **RESIST** está diseñando nuevos enfoques para supervisar y configurar la infraestructura de la cadena de bloques (red, nubes, etc.) con el fin de mejorar o garantizar el rendimiento y la seguridad. Además, el equipo también investiga la gestión basada en blockchain para abordar la evolución del ecosistema de Internet.
↗ El equipo **SPECFUN** realiza pruebas formales de teorías matemáticas, así como análisis formales del entorno de Ethereum y otros modelos de ejecución habilitados para blockchain.
↗ El equipo **TOCCATA** promueve la especificación formal y la prueba asistida por computadora en el desarrollo de software que requiere una alta garantía de su seguridad y corrección.
↗ El equipo **VERIDIS** estudia técnicas de verificación mecanizada aplicadas a algoritmos y sistemas concurrentes y distribuidos. Se interesa por formulaciones precisas de los problemas algorítmicos que surgen en el contexto de las cadenas de bloques y los correspondientes problemas de verificación.

6.2 Sistemas críticos y ciberfísicos

Los sistemas críticos son sistemas que requieren una fiabilidad muy alta, ya que un fallo podría tener consecuencias extremadamente perjudiciales, como poner en peligro vidas humanas, causar graves daños a una infraestructura o provocar importantes pérdidas económicas. Sin embargo, la mayoría de los sistemas críticos para la seguridad lo son también respecto a su propia seguridad y deben ser resistentes a los ciberataques.

Un concepto relacionado es el de sistema ciberfísico (Cyber Physical System o CPS, por sus siglas en inglés). Wikipedia define un CPS como *un mecanismo controlado o monitorizado por algoritmos basados en computación, estrechamente integrados con Internet y sus usuarios. En los sistemas ciberfísicos, los componentes físicos y de software están profundamente entrelazados, donde cada elemento opera en diferentes escalas espaciales y temporales, exhibiendo múltiples comportamientos, e interaccionando entre ellos de innumerables formas que cambian con el contexto.* Los CPS suelen formar parte de sistemas críticos, que deben ser seguros. Compuestos por muchos subsistemas interconectados, funcionando de forma diferente de comunicación a diferentes escalas, ofrecen una gran superficie de ataque. Además, algunos de ellos son objetivos con riesgo porque un ataque exitoso, como el cierre de parte de una infraestructura estatal, puede tener un enorme impacto económico y político. Por lo tanto, la ciberseguridad

se ha convertido en una cuestión esencial para los CPS.

Aunque la ciberseguridad de los CPS puede reutilizar los enfoques, métodos y técnicas tradicionales para asegurar los sistemas y las redes, también habrá que desarrollar nuevos enfoques para hacer frente a la dinamicidad y para analizar y garantizar su seguridad en este contexto. Es interesante que éstos se inspiren en la seguridad reactiva, donde la supervisión y la reacción ante situaciones anómalas ocupan un lugar más importante, y es probable que el aprendizaje automático desempeñe un papel cada vez más importante que, por supuesto, dará lugar a sus propios problemas de seguridad.

Seguridad de los sistemas integrados (fuzzin) – © Inria / photo C. Morel

[Nota] *CPS frente a los* **sistemas** *integrados*

Los CPS generalizan el concepto mucho más antiguo de los sistemas integrados en muchos sentidos. Un sistema integrado también implica una interfaz entre el mundo digital y el físico y hace hincapié en las cuestiones temporales, pero se centra en los componentes individuales en lugar de en todo el sistema y en las interacciones entre sus numerosos componentes y, por supuesto, también en las personas. Por ejemplo, el ABS (sistema antibloqueo de frenos) es un sistema integrado y, por tanto, también un CPS, mientras que el controlador de una central eléctrica es un CPS en el que intervienen muchos otros subsistemas. La gran mayoría de los dispositivos informáticos que se utilizan hoy en día son en realidad partes de sistemas integrados o CPS.

6.2.1 Seguridad de la Internet de las cosas (IoT)

[Resumen]

La revolución de la IoT está impulsando la extensión de la Internet a gran velocidad y está cambiando la forma en que el mundo interactúa con los dispositivos físicos. Las consideraciones de seguridad y privacidad ponen en juego esta revolución. Las aplicaciones de la IoT no son una manifestación más de los sistemas distribuidos conocidos hasta ahora: sus particularidades, en particular la naturaleza de los dispositivos con recursos limitados, plantean nuevos retos para la seguridad que deben abordarse y las soluciones de seguridad convencionales no siempre son aplicables. Hay que explorar varias líneas de investigación, en particular: sistemas operativos seguros y capacidades de actualización de firmware, criptografía ligera y asistida por hardware, seguridad y privacidad de las tecnologías inalámbricas, políticas de seguridad y privacidad específicas para IoT, métodos de detección de intrusiones en la IoT, lenguajes de programación y compiladores seguros para aplicaciones de la IoT, y protocolos de autenticación específicos.

La Internet de las Cosas (IoT) es la red de dispositivos físicos integrados, equipados con sensores, activadores, procesamiento, almacenamiento y conectividad que permite a estas cosas conectarse e intercambiar datos (Wikipedia). Las aplicaciones de la IoT pertenecen a una gran variedad de dominios, desde la salud de los pacientes (por ejemplo, los marcapasos inteligentes), dispositivos auto-cuantificados (por ejemplo, relojes de seguimiento de la actividad), hasta los electrodomésticos (por ejemplo, termostatos, enchufes o ampolletas inteligentes), los autos (por ejemplo, sistemas de emergencia de vehículo a vehículo) y los sistemas industriales (por ejemplo, SCADA). Dado que la IoT permite una integración más directa del mundo físico en los sistemas informáticos, ha sido calificada[113] como la Cuarta Revolución Industrial.

Sin embargo, las preocupaciones sobre la seguridad de la IoT, o la falta de ella, y la privacidad, ya que se intercambian datos personales, a veces sensibles, ponen en juego esta revolución. Por ejemplo, la red de bots Mirai (apartado 1.3) utilizó un gran número de dispositivos de consumo (incluidas cámaras IP vulnerables) para lanzar ataques masivos sobre sistemas de discos operativos (DoS) dispersos. Del mismo modo, la privacidad de los usuarios está en riesgo cuando las empresas tratan la seguridad como algo secundario (como fue el caso de algunos juguetes sexuales conectados a Internet[114]).

Una primera particularidad de la IoT es la naturaleza limitada de los recursos de los dispositivos. Los microcontroladores de los dispositivos IoT tienen una arquitectura muy diferente a la de un PC típico en cuanto a capacidades de almacenamiento y CPU[115]. Algunos dispositivos o componentes de la IoT estarán estrictamente acotados no solo por el bajo consumo de energía, sino también por la limitación de recursos. Estos dispositivos

113. Fuente: Foro Económico Mundial
114. https://blog.trendmicro.com/penetration-testing-researchers-successfully-hack-a-vibrator/
115. Por ejemplo, la memoria RAM y el almacenamiento Flash son 106 veces más pequeños en un microcontrolador Arduino Uno y su potencia de procesamiento es de unos 16 MIPS (millones de instrucciones por segundo), frente a unos 100.000 MIPS de los procesadores de sobremesa recientes.

de consumo y costo ultra bajos funcionan con baterías de capacidad limitada o incluso recogen energía del entorno (por ejemplo, luz, calor o vibración).

Estas particularidades condujeron al diseño de sistemas operativos específicos para manejar distintas arquitecturas integradas. Entre los sistemas operativos de la IoT más conocidos se encuentran Contiki[116] (iniciado en el Instituto Sueco de Ciencias Informáticas), Mbed OS[117] (ARM), RIOT[118] (iniciado por Inria, la Universidad Libre de Berlín y la Universidad de Ciencias Aplicadas de Hamburgo) y Zephyr[119] (Wind River Systems, Linux Foundation, Intel, NXP Semiconductors, etc.).

Otra singularidad es la naturaleza heterogénea de una aplicación global de IoT. Al cubrir el espectro desde los microprocesadores hasta la nube, una aplicación IoT incluye potencialmente código que se ejecuta en clientes de la Web, en servidores, así como en dispositivos integrados. Las aplicaciones IoT deben manejar una gran variedad de eventos asíncronos: consultas a servicios lejanos, respuestas, tiempos de espera y errores. A su vez, cada evento puede lanzar cálculos que desencadenen una cascada de nuevos eventos.

Llevar la seguridad y la privacidad a la IoT es un reto, ya que la superficie de ataque es significativa y las soluciones de seguridad tradicionales no siempre son aplicables. Entre las líneas de investigación que deben abordarse, podemos mencionar las siguientes:

POSIBILIDADES SEGURAS DE ACTUALIZACIÓN DE SOFTWARE

El despliegue seguro de las actualizaciones de software en los dispositivos IoT, y en particular del propio firmware, es una característica obligatoria de cualquier sistema operativo seguro. Por desgracia, esta característica clave se ve a menudo comprometida por el deseo de mantener la complejidad del dispositivo y el consumo de energía lo más bajo posible, o simplemente por malas prácticas. Además de los trabajos académicos, el grupo de trabajo IETF SUIT[120] (Software Updates for Internet of Things) investiga esta cuestión, centrándose en una arquitectura ajena a las tecnologías y protocolos de comunicación (por ejemplo, CoAP o HTTP).

CRIPTOGRAFÍA LIGERA Y CRIPTOGRAFÍA ASISTIDA POR HARDWARE

En cuanto a los bloques de construcción criptográficos, las limitaciones de los dispositivos integrados pueden requerir el uso de primitivas criptográficas ligeras. Este tema se trata en el apartado 3.1.3.

Los dispositivos IoT también se beneficiarán del diseño de arquitecturas de hardware específicas y de la optimización del software de criptografía ligera, incluida la protección contra los ataques de canal lateral y la inyección de fallos. Estos objetivos requieren una plataforma eficiente basada en:

- aceleradores de hardware (criptoprocesadores) para funciones de seguridad (por

116. http://contiki-os.org/
117. https://www.mbed.com/en/platform/mbed-os/
118. https://www.riot-os.org/
119. https://www.zephyrproject.org/
120. https://datatracker.ietf.org/wg/suit

ejemplo, criptografía simétrica y asimétrica, hashing, autenticación, firma o generadores de números aleatorios), con un enfoque específico en la eficiencia energética y el consumo ultra bajo;

- diseños exclusivos de criptoprocesadores, como la randomización, que protegen contra los ataques;
- optimizaciones del compilador para criptoprocesadores con recursos limitados;
- traducción binaria dinámica acelerada por hardware (DBT) como medio para mejorar la protección del software;
- y nuevas técnicas para la protección eficaz del hardware contra los ataques de canal lateral y la inyección de fallos, tanto en el software como en el hardware.

El objetivo global es presentar un criptoprocesador energéticamente eficiente con una arquitectura de conjunto de instrucciones que se pueda configurar en tiempo real.

SEGURIDAD DE TECNOLOGÍAS INALÁMBRICAS

Las comunicaciones de la IoT se apoyan tanto en tecnologías inalámbricas estándar (por ejemplo, wifi y Bluetooth) como en tecnologías de baja potencia (por ejemplo, Bluetooth Low Energy (BLE) y Zigbee), así como en redes de baja potencia y gran superficie en bandas sin licencia (por ejemplo, LoRa[121] y SigFox[122]) y en bandas con licencia (la prometedora 5G que también pretende conectar dispositivos de la IoT). Se está trabajando en los riesgos de seguridad y privacidad (apartado 5.3.6) asociados a estas tecnologías de comunicación.

POLÍTICAS DE SEGURIDAD Y PRIVACIDAD ESPECÍFICAS PARA LA IOT

Otro reto es el diseño de políticas de seguridad especiales para la IoT que impidan ataques que exploten la exposición del sistema a peligros físicos o que tengan consecuencias físicas.

MÉTODOS DE DETECCIÓN DE INTRUSOS EN LA IOT

Otro reto para la seguridad de la IoT es el diseño de técnicas de detección de intrusiones en la IoT para detener el malware y otros ataques, por ejemplo, basados en las actividades de los dispositivos en las redes de la IoT (apartado 4.4).

LENGUAJES DE PROGRAMACIÓN Y COMPILADORES SEGUROS PARA APLICACIONES IOT

La programación de una aplicación IoT es compleja debido a la mezcla de tecnologías implicadas y a los lenguajes utilizados, y tener en cuenta la seguridad y la privacidad añade otro nivel de complejidad. Existen varios marcos destinados a simplificar el desarrollo de la IoT, pero no siempre están diseñados teniendo en cuenta la seguridad y pueden presentar graves riesgos. Por lo tanto, es necesario desarrollar un lenguaje de programación y un marco de trabajo seguros y que preserven la privacidad para

121. https://en.wikipedia.org/wiki/LoRa
122. https://en.wikipedia.org/wiki/Sigfox

la IoT. La verificación formal del software de la IoT, empezando por su biblioteca criptográfica, es otra posible línea de investigación.

PROTOCOLOS DE AUTENTICACIÓN ESPECÍFICOS

Por último, la combinación de dispositivos con recursos limitados y la heterogeneidad de la IoT requiere nuevos protocolos de autenticación para proporcionar una forma sencilla, segura y respetuosa con la privacidad que permitan autenticar sin problemas a las personas y sus dispositivos inteligentes en línea.

[Equipos Inria] Seguridad de la Internet de las Cosas

↗ El equipo **ANTIQUE** está trabajando en los análisis basados en la interpretación abstracta para los programas de la IoT.

↗ El equipo **CAIRN** aborda la cuestión de mejorar la eficiencia energética mediante el uso de aceleradores de hardware flexibles. Propone basarse en un hardware transformable, cuya estructura puede reconfigurarse de forma eficiente en tiempo real de ejecución, por ejemplo, para hacer que el hardware se adapte a la aplicación que ejecuta y adaptar el paralelismo del acelerador para el rendimiento o el uso de recursos. El equipo también explora la traducción binaria dinámica acelerada por hardware.

↗ El equipo **CELTIQUE** está trabajando en la formalización de la semántica de Hop.js en Coq, así como en la verificación formal de los análisis de programas para la IoT.

↗ El equipo **FUN** está trabajando en la verificación del sistema Contiki, así como en técnicas de detección de intrusiones para la IoT.

↗ El equipo **GRACE** participó en el diseño de varias primitivas criptográficas específicas para microcontroladores, en particular qDSA, un esquema de firma para la IoT, ahora incluido en el sistema operativo RIOT.

↗ El equipo **INDES** está trabajando en el desarrollo de un compilador seguro y en la aplicación de políticas de privacidad para Hop.js, un lenguaje de programación para la IoT.

↗ El equipo **INFINE** lidera el desarrollo del sistema operativo de código abierto para IoT RIOT[a], junto con la Universidad Libre de Berlín y la Universidad de Ciencias Aplicadas de Hamburgo. Se está trabajando en el soporte de instalaciones de actualización de firmware seguras para RIOT.

↗ El equipo **KAIROS** está trabajando en técnicas de co-diseño basadas en modelos formales para los temas de tiempo real y seguridad de la IoT.

↗ El equipo **PETRUS** está trabajando en estrategias adaptadas para almacenar datos en los objetos de la IoT en lugar de enviar los datos a los servidores.

↗ El equipo **PRIVATICS** está trabajando en consideraciones de privacidad y sistemas de preservación de la privacidad para la IoT.

↗ El equipo **PROSECCO** está construyendo HACL*, una biblioteca criptográfica verificada que ya está incluida en RIOT y se está trabajando en una versión ligera para dispositivos IoT con recursos limitados.

a. https://www.riot-os.org/

[Desafío de investigación 10] Asegurar la Internet de las Cosas (IoT)
La seguridad en la IoT es un reto importante: los ataques son todavía relativamente fáciles (muchos dispositivos no se han diseñado teniendo en cuenta la seguridad), invasivos (por ejemplo, omnipresentes en nuestras vidas) y con un impacto potencialmente importante debido al factor de multiplicación que permite el gran número de dispositivos disponibles y a las implicaciones directas que algunos de ellos tienen en el mundo físico (por ejemplo, los automóviles conectados). Las líneas de investigación son múltiples, como por ejemplo, la capacidad de actualizar de forma segura el software de los dispositivos integrados, el diseño de bases primitivas criptográficas ligeras adaptadas a los recursos limitados, el análisis de la seguridad de las nuevas tecnologías inalámbricas de baja potencia y gran área, la detección y mitigación de intrusiones o dispositivos que se comportan erróneamente, y la necesidad de marcos, protocolos y sistemas operativos seguros por diseño (secure-by-design).

6.2.2 Seguridad de los sistemas industriales

[Resumen]
La ciberseguridad de los sistemas industriales es un tema emergente y los recientes ciberataques a sistemas industriales demuestran que el problema está sin resolver. Una de las principales dificultades para tratar con los sistemas industriales es que no cumplen el principio de seguridad por diseño: como no fueron pensados en un principio para ser expuestos en Internet, los protocolos utilizados no son seguros; a veces las especificaciones no están disponibles públicamente; los dispositivos de detección de intrusos y los firewalls de uso general no manejan los protocolos industriales. Los dispositivos finales se construyen con procesadores lentos incapaces de utilizar los protocolos criptográficos estándar y, por tanto, requieren otros específicos.

Los sistemas de control industrial (ICS) o los sistemas de control y automatización industrial (IACS) incluyen una gran variedad de sistemas digitales industriales de adquisición de datos, control y supervisión junto con sus redes de comunicación subyacentes. Actualmente se utilizan dos arquitecturas de sistemas: DCS (Distributed Control System, o Sistema de Control descentralizado) y SCADA (Supervisory Control and Data Acquisition, o control de supervisión y obtención de datos), aunque SCADA tiende a sustituir a DCS.

La ciberseguridad de los sistemas de control industrial es un campo de aplicación relativamente reciente. Las comunicaciones en los sistemas industriales tradicionales se realizaban a través de redes de pequeño tamaño que utilizaban protocolos propios. Estaban sometidas a duras restricciones de tiempo real y no estaban interconectadas con sistemas informáticos o con Internet. Por lo tanto, la ciberseguridad tradicional en los sistemas industriales se lograba mediante aislamiento y ocultación, utilizando

Comprender los ataques a los sistemas de control industrial – © Inria / Photo C. Morel

protocolos propios. Los sistemas industriales modernos y de gran envergadura, como las redes eléctricas, las centrales nucleares o las presas hidráulicas, tienen necesidades de optimización de control global que exigen la interconexión con aplicaciones de control de supervisión, sistemas de bases de datos distribuidos y, en consecuencia, comunicación de largo alcance, formatos de intercambio de datos estándar e interoperabilidad. El paradigma de comunicación primordial en los sistemas industriales modernos es la convergencia de la Tecnología de la Información (TI) y la Tecnología de Operación (TO), o dicho de otro modo, la interconexión entre las redes TCP/IP y los protocolos propios en tiempo real. La contrapartida de la penetración de las TI en el ámbito de las TO es que éstas están ahora expuestas a los ciberataques del mismo modo que las TI. Además, como la TO sigue dependiendo en gran medida de los protocolos heredados, es aún más vulnerable. Sin embargo, durante casi una década (1995-2003), la ciberseguridad de los sistemas industriales se ignoró en gran medida.

El acontecimiento que inició la investigación sobre la ciberseguridad de los sistemas industriales fue el apagón del noreste de Estados Unidos en 2003. Aunque no fue el resultado de un ciberataque (para dejar constancia, el apagón se debió a un sensor mal calibrado y a una cadena de fallos informáticos), este acontecimiento demostró que una falsa inyección de datos puede explotar varias vulnerabilidades para cerrar 256 centrales nucleares y dejar sin luz a 55 millones de personas en Estados Unidos y Canadá. Hasta 2010 se avanzó poco y no se demostró formalmente la realidad de la amenaza. Tras la aparición de Stuxnet en 2010-2011 se iniciaron importantes programas de investigación. Otros acontecimientos, como la cadena de ataques a la red eléctrica ucraniana en 2015-2016, aumentaron la importancia de programas de ciberseguridad industrial.

Las soluciones de ciberseguridad informática no se aplican directamente, debido a las especificidades de los sistemas industriales. A continuación se presenta una lista de puntos débiles de la comunicación ICS:

· Protocolos inseguros y no asegurables: los protocolos heredados no fueron diseñados para la seguridad; además, también están pensados para ser utilizados con procesadores de muy baja velocidad, lo que descarta el uso de protocolos criptográficos.

· Especificaciones no disponibles: algunos protocolos heredados siguen siendo privados y no se divulgan las especificaciones completas.

· Muchas versiones: algunos protocolos heredados están pensados para ser extendidos por desarrolladores con sus propios mensajes.

· Gran superficie de ataque: los dispositivos terminales (por ejemplo, los controladores lógicos programables) suelen actuar como pasarelas de red entre varias redes heredadas.

· Ataques orientados al proceso: la legitimidad de un paquete de red depende del estado del sistema físico subyacente o de la frecuencia del paquete. Por ejemplo, abrir un grifo de alimentación de un tanque será inofensivo si el tanque está en un nivel intermedio, pero puede dañar la planta si el tanque está lleno. Del mismo modo, Stuxnet sólo modificó el valor de un comando de control que, por lo demás, es legítimo. De todas formas, arrancar y parar repetidamente un accionador puede acabar dañándolo.

La detección de intrusos orientada a procesos es un tema importante en el ámbito de la ciberseguridad de los ICS.

[Equipos Inria] Seguridad de los sistemas industriales

↗ Los equipos **CIDRE** y **CTRL-A** desarrollaron un enfoque basado en la supervisión de las especificaciones de los procesos. Las especificaciones de los procesos son propiedades de seguridad expresadas en lógica temporal lineal (LTL). Se extraen automáticamente de los rastros de ejecución y se supervisan mediante técnicas de verificación en tiempo real. Se utiliza la correlación de alertas entre los monitores de procesos, un sistema de detección de intrusos de red (IDS) de lista de aprobación y un IDS de patrones para reducir el número de falsos positivos. El enfoque se valida en un banco de pruebas de ICS.

↗ El equipo **CTRL-A** también ha desarrollado un banco de pruebas y demostradores de ataque/defensa en sistemas industriales y redes inteligentes. En cuanto a la investigación de la vulnerabilidad de los protocolos, el equipo **CTRL-A** trabajó en una vulnerabilidad del protocolo en tiempo real en las redes IEC 61850, demostrando el ataque, y propuso un módulo IDS en BRO.

[Desafío de investigación 11] Sistemas industriales seguros

Los sistemas industriales se basan cada vez más en mecanismos de software que pueden ser atacados. Por ello, su seguridad se ha convertido en un problema importante,

sobre todo porque las consecuencias de un ataque contra estos sistemas pueden ser dramáticas. Aunque los enfoques de seguridad tradicionales parecen aplicables al caso de los sistemas industriales, sus especificidades exigen revisar los mecanismos de seguridad tradicionales para adaptarlos a este nuevo contexto. En particular, los protocolos de comunicación utilizados en este ámbito no pueden modificarse de la noche a la mañana. Debe haber una transición durante la cual las comunicaciones heredadas deben integrarse en protocolos seguros. Además, suele ser necesario controlar el sistema en tiempo real. Por tanto, la seguridad también debe ser aplicable en tiempo real. Por último, a menudo es imposible modificar los dispositivos industriales. Por lo tanto, no se pueden utilizar mecanismos de seguridad preventivos. Por todo ello, es obligatorio utilizar la seguridad reactiva y, por lo tanto, es muy importante estudiar cómo se pueden desplegar mecanismos de detección de ataques eficaces en este contexto.

6.3 Áreas críticas de aplicación

Algunos ámbitos de aplicación son especialmente sensibles desde el punto de vista de la seguridad. La medicina es un ejemplo, ya que los datos médicos son extremadamente sensibles y los ataques a los equipos e implantes médicos podrían poner en peligro la vida de las personas.

Por último, el aprendizaje automático se ha hecho muy popular recientemente en muchos ámbitos. Aunque no es un área de aplicación en sí misma, es un caso de estudio interesante para la seguridad: las técnicas de aprendizaje automático pueden ser engañadas por los llamados "ejemplos adversarios" y presentan riesgos para la privacidad cuando se entrenan con datos sensibles.

6.3.1 Medicina

[Resumen]
La medicina está siendo transformada significativamente con la revolución digital y, por lo tanto, está cada vez más expuesta a las amenazas de ciberseguridad. Una de ellas es la filtración de datos médicos sensibles. Dado que estos datos son extremadamente útiles para la investigación, el equilibrio entre privacidad y utilidad es de especial importancia. Los aparatos médicos, como los robots de asistencia quirúrgica, están cada vez más conectados. Del mismo modo, los implantes médicos ofrecen conexiones inalámbricas para evitar cirugías innecesarias, lo que aumenta la superficie de ataque de manera similar a la de las CPS y la IoT.

La medicina, al igual que muchos otros ámbitos, se está transformando significativamente por la revolución digital y, por lo tanto, también está cada vez

más expuesta a las amenazas de ciberseguridad. Aunque la ciberseguridad en la medicina no es muy diferente de la ciberseguridad en otros ámbitos, muchos aspectos de la ciberseguridad son, de hecho, más delicados en la medicina, ya que esta implica, en última instancia, vidas humanas, para las que no hay precio y, por tanto, no hay riesgo aceptable. Esto se refiere, por supuesto, cuando un ataque podría poner en peligro nuestras vidas, pero también cuando podría filtrar nuestros historiales médicos que pertenecen a la clase de Información Personal Sensible y, por tanto, están sujetos a una normativa muy protectora (véase el Capítulo 5).

Por ello, resulta instructivo considerar la medicina como un dominio de aplicación. Aunque la privacidad ha sido durante mucho tiempo una seria preocupación en la medicina, sigue siendo una cuestión difícil; a menudo se subestiman otros aspectos de la ciberseguridad que podrían poner en peligro nuestras vidas debido a acciones maliciosas.

Efectivamente, aunque la seguridad del paciente siempre ha formado parte de la cultura de la medicina, en la que se toman todas las precauciones posibles para salvar nuestras vidas, la seguridad no suele ser una gran preocupación, porque los hospitales y centros asistenciales se consideran normalmente un santuario en el que no hay lugar para la delincuencia. Esto puede estar cambiando con la digitalización de la medicina, ya que los ataques pueden realizarse a distancia, desde fuera del hospital.

Existen principalmente tres tipos de amenazas: filtraciones de privacidad, acciones maliciosas en los aparatos médicos y ataques a los implantes.

FILTRACIONES DE PRIVACIDAD

Nuestra preocupación –y por tanto nuestra concientización– por las filtraciones de privacidad es mucho mayor que por las acciones maliciosas, probablemente porque incluso antes de la digitalización, siempre hemos tratado la información sanitaria como algo muy privado y personal. Sin embargo, los datos médicos están cada vez más centralizados y se almacenan durante periodos muy largos, por lo que, a pesar de nuestra concienciación, este sigue siendo un problema crítico que aún no ha encontrado un compromiso aceptable.

Aunque los datos médicos son muy sensibles y no deberían filtrarse, la creciente cantidad de datos médicos a lo largo de la vida de una persona y en una población constituye una información valiosa para la investigación médica. En efecto, un estudio médico basado en decenas o centenas de pacientes dista mucho de ser estadísticamente suficiente. Pronto tendremos la posibilidad de analizar no solo miles o incluso millones de registros, sino también registros que contienen información extremadamente rica para la que las técnicas de aprendizaje automático podrían revelar interesantes correlaciones entre patologías y factores fisiobiológicos. Aunque mantener los datos médicos privados y descentralizados sería una buena respuesta a la preservación de la privacidad, podría ser un obstáculo para

el progreso médico, dado el considerable impacto potencial que el acceso a estos datos podría tener para los descubrimientos médicos.

APARATOS MÉDICOS

La sanidad es un importante usuario de equipos digitales. Si durante mucho tiempo ha sido el caso de los equipos de imagen médica, recientemente se ha extendido a todo tipo de monitorización, así como a los asistentes quirúrgicos robóticos. Todos estos sistemas están, por supuesto, interconectados, lo que aumenta la superficie de ataque. No hay particularidades que mencionar sobre los aparatos, salvo que en general no se presta la suficiente atención a las amenazas de ciberseguridad y a veces ni siquiera a la seguridad de estos dispositivos

IMPLANTES

Más preocupantes son los implantes médicos (que van desde los marcapasos, los implantes de insulina o los implantes de auditivos hasta los corazones artificiales) que hoy en día pueden controlarse a distancia, permitiendo el ajuste de los parámetros sin una intervención quirúrgica. Esto los expone automáticamente a ataques de ciberseguridad como en el caso de los sistemas ciberfísicos (véase el apartado 6.2). Sin embargo, estos implantes suelen estar sometidos a importantes limitaciones de tamaño y consumo de energía, lo que hace que las técnicas normales de encriptación sean inaplicables, como suele ocurrir con los dispositivos IoT. Los fabricantes eligen entonces soluciones ad hoc que probablemente sean insuficientemente seguras.

[Inria teams] Medicina

↗ El trabajo del equipo **PETRUS** sobre la nube privada (apartado 4.2.2) fue motivado originalmente como una forma de descentralizar los datos médicos, permitiendo que todo el mundo tenga su historial médico personal y mantenga el control sobre qué, cómo y por quién se puede acceder a la información cuando sea necesario.

↗ Los equipos **PRIVATICS** y **TYREX** trabajan conjuntamente en la mejora de los vínculos entre privacidad y utilidad (apartado 5.1.1), las técnicas de anonimización (apartado 5.2.2) y la privacidad del aprendizaje automático (apartado 6.3.3).

6.3.2 Robótica y vehículos autónomos conectados

[Resumen]

Podemos dividir la robótica en cuatro grandes áreas de aplicación: robots autónomos (por ejemplo, robots de fábrica, robots para asistir a los ancianos), robots operados a distancia (por ejemplo, robots quirúrgicos, drones), transporte basado en la robótica (por ejemplo, vehículos terrestres con o sin personas dentro), y grandes sistemas de sistemas robóticos (por ejemplo, ciudades inteligentes). En Inria, la mayor parte de

la investigación en torno a la ciberseguridad y la robótica se centra en los vehículos autónomos conectados (CAV). Los coches conectados se comunican entre sí y también con objetos externos. Por cuestiones de tradición, las arquitecturas no aíslan las partes críticas de las no críticas. Esto hace que los CAV sean vulnerables tanto a las amenazas internas como a las externas. Las amenazas internas suelen deberse a fallos del software o a canales no seguros. Las externas se deben a ataques remotos que explotan los canales de comunicación. Las líneas actuales para evitar estos ataques tienden a aislar los subsistemas críticos de los no críticos.

Los sistemas de robots conectados son el reflejo de los sistemas ciberfísicos. Los primeros ejemplos se encuentran en fábricas y lugares peligrosos (por ejemplo, plantas nucleares o químicas). El acercamiento entre humanos y robots es cada vez más estrecho. De forma esquemática, se pueden distinguir cuatro áreas de aplicación, que se entrecruzan parcialmente: los robots autónomos (por ejemplo, robots de fábrica, robots para asistir a los ancianos), los robots operados a distancia (por ejemplo, robots quirúrgicos, drones), el transporte basado en la robótica (por ejemplo, vehículos terrestres con o sin personas dentro) y los grandes sistemas de sistemas robóticos (por ejemplo, las ciudades inteligentes). Por supuesto, existe una intersección no vacía entre los sistemas de robots conectados y la IoT. Las ciberamenazas pueden conducir a resultados no deseados, posiblemente catastróficos e irreversibles, en el ciberespacio y en el espacio físico.

Un reciente estudio[123] de TrendMicro y la Universidad Politécnica de Milán ha demostrado que los robots industriales conectados a Internet son vulnerables y no están protegidos. Este problema de seguridad también se conoce en el caso de los robots quirúrgicos accionados a distancia[124]. Los especialistas en robótica están prestando cada vez más atención a los problemas de ciberseguridad, debido a la aparición de hardware y SoC de bajo costo que están cambiando el aspecto económico del ámbito. La posibilidad de contar con numerosos "nodos inteligentes" asequibles (con servicios de ciberseguridad de bajo consumo y computacionalmente eficientes implementados en ellos) dentro de las redes de robots o IoT abre nuevas perspectivas.

En la actualidad, dentro de Inria, el ámbito relacionado con la robótica en el que se abordan las cuestiones de ciberseguridad es el de los vehículos autónomos conectados (CAV), en particular los autos –véase el libro blanco dedicado a este tema–[125]. Además de la robótica a bordo, los CAV estarán equipados con dispositivos de radio que permitirán el intercambio de datos con otros vehículos (comunicaciones vehículo

123. https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/rogue-robots-testing-industrial-robot-security
124. http://www.washington.edu/news/2015/05/07/uw-researchers-hack-a-teleoperated-surgical-robot-to-reveal-security
125. Véhicules autonomes et connectés - les défis actuels et les voies de recherche. Disponible en https://www.inria.fr/institut/strategie/vehicules-autonomes-et-connectes

a vehículo, o V2V), así como el acceso móvil a servicios basados en Internet (comunicaciones vehículo a todo, o V2X). Además, los CAV estarán equipados con sensores internos capaces de controlar a los pasajeros, en particular su somnolencia. Sin embargo, en los vehículos actuales, todos los dispositivos comparten el mismo bus de comunicación y la mayoría de los actuadores (por ejemplo, la dirección o los frenos) son accesibles a través del puerto de diagnóstico de a bordo. Además, los sistemas de a bordo actuales son monolíticos. Por lo tanto, los CAV actuales son vulnerables a las ciberamenazas tanto internas como externas.

PROBLEMAS DE CIBERSEGURIDAD INTERNA

Los problemas internos se deben a atacantes nativos (software defectuoso, puertas traseras) y a intrusiones a través de canales no seguros (virus, malware). Las intrusiones pueden hacer que un vehículo, por lo demás fiable, se comporte de forma maliciosa. Esto se ha ilustrado con la toma de control de un Jeep Cherokee mediante un troyano colocado en un CD de MP3.

Las cámaras y los diversos sensores instalados en los CAV plantean problemas de privacidad específicos. ¿A quién pertenecen los datos recogidos? ¿Qué ocurre si los datos son robados? Además, la lógica subyacente según la cual un pasajero debería poder retomar el control siempre que sea necesario no está ampliamente reconocida, ni aceptada, ya que los retrasos de la reacción humana suelen considerarse demasiado elevados.

PROBLEMAS DE CIBERSEGURIDAD EXTERNOS

Los ataques externos pueden realizarse a distancia a través de las comunicaciones V2X. Los problemas de ciberseguridad que plantean las tecnologías V2X son similares a los que surgen en cualquier servicio de red. Y dado que las comunicaciones V2X dependen de los relés intermedios (por ejemplo, los dispositivos en carretera y los nodos de la red de telecomunicaciones), sin una protección específica, favorecen los ataques de suplantación de identidad y, por tanto, el falseamiento, la supresión o falsificación identitaria de mensajes.

En Francia, las autoridades estatales y las compañías de seguros exigen la autenticación y no repudio para la correcta identificación de las responsabilidades (por ejemplo, en caso de accidente). Por ello, en cada vehículo se almacena un certificado emitido por el Estado en un dispositivo a prueba de manipulaciones. Para mejorar la seguridad de los CAV, se está estudiando la posibilidad de emitir balizas periódicas. Esto implica la transmisión de la posición, la velocidad y la dirección sin cifrar varias veces por segundo. Las identidades reales se enmascararían utilizando el seudónimo, es decir, el anonimato reversible basado en la criptografía asimétrica[126]. Desgraciadamente, se ha demostrado que los cambios frecuentes de seudónimos no impiden la vinculación de las trayectorias de los

126. https://research.utwente.nl/en/publications/pseudonym-schemes-in-vehicular-networks-a-survey

CAV. Esto supone un verdadero problema de privacidad y la emisión de balizas ha quedado en entredicho.

Para evitar los ataques anteriores, los científicos y los fabricantes están definiendo arquitecturas compartimentadas que constan de dos subsistemas. Un subsistema crítico de seguridad se encarga de la mensajería crítica V2V, el procesamiento y la conducción cooperativa. Un subsistema no crítico para la seguridad, aislado de la parte crítica, se encarga de los servicios basados en las comunicaciones V2X (por ejemplo, infoentretenimiento, tráfico y condiciones de la carretera, plazas de aparcamiento libres, mensajería personal). El subsistema crítico de seguridad estará dotado de la capacidad de inspeccionar los mensajes V2X recibidos por el subsistema no crítico de seguridad, antes de importarlos para su posterior procesamiento.

[Equipos Inria] Vehículos autónomos conectados

↗ El equipo **CIDRE** trabaja actualmente en la definición de un sistema de detección de intrusos (IDS) que podría integrarse en la arquitectura de los vehículos de próxima generación. En el contexto de los CAV deben explorarse varios escenarios (IDS basado en la red o en el host, anomalía o detección).

↗ El equipo **RITS** está investigando los subsistemas y protocolos críticos para la seguridad de las comunicaciones de los vehículos. El objetivo es permitir la prevención de ataques, su detección inmediata, y que las escuchas y el seguimiento sean inviables e inservibles. Además del seudónimo, se ha demostrado que el anonimato (ofuscación no reversible de las identidades) de los emisores de mensajes es factible mediante la formación espontánea de subredes vehiculares ad hoc de confianza. Los sistemas de a bordo disponen de una opción de "modo oculto" (sin envío de mensajes V2X), para aumentar la privacidad.

6.3.3 Tecnologías basadas en el aprendizaje automático

[Resumen]

El aprendizaje automático se utiliza en un número cada vez mayor de aplicaciones, como el comercio electrónico y los sistemas de recomendación, los mecanismos avanzados de traducción de idiomas, las herramientas de filtrado de spam o parental, así como los coches auto-conducidos y los sistemas ciberfísicos en general. El impacto más profundo hasta el momento se produce sin duda en el reconocimiento del habla y de la imagen, pero otros ámbitos también se ven afectados de forma significativa. Las técnicas de aprendizaje automático sufren dos amenazas principales en relación con la ciberseguridad. La primera, denominada aprendizaje automático adverso, consiste en añadir ruido cuidadosamente diseñado (apenas visible para el ojo humano) a una imagen, lo que conduce a una clasificación errónea. La segunda está relacionada con

la privacidad y consiste en extraer información sobre los datos de aprendizaje de una red entrenada.

Las tecnologías basadas en el aprendizaje automático (ML) constituyen ahora la columna vertebral de un número cada vez mayor de organizaciones y servicios. Entre ellos se encuentran los sistemas de comercio electrónico y de recomendación, los mecanismos avanzados de traducción de idiomas, las herramientas de filtrado de spam o parental, así como los autos que se conducen solos. Millones de usuarios interactúan a diario con estos sistemas, de forma transparente, incluso sin darse cuenta.

Hay un ámbito en el que las redes neuronales convolucionales y las estrategias de aprendizaje profundo han tenido un gran impacto: la visión por computadora. Los problemas tradicionales, como la clasificación precisa de imágenes o la detección y el etiquetado de objetos de grano fino en imágenes, han experimentado recientemente un enorme progreso, con enfoques de última generación que superan con creces los métodos más antiguos e incluso superan las capacidades humanas para entornos específicos.

APRENDIZAJE AUTOMÁTICO ADVERSO

Investigaciones recientes han demostrado que estos enfoques de aprendizaje automático pueden subvertirse cuando se añade a los datos de entrada una pequeña cantidad de ruido adverso cuidadosamente diseñado e imperceptible. Esta imagen distorsionada, denominada ejemplo adverso, suele ser clasificada erróneamente, aunque la perturbación apenas sea visible para el ojo humano. La perturbación adversa se aplica sorprendentemente bien a través de diferentes imágenes de entrada, clasificadores y modelos, incluso cuando se entrena en diversos conjuntos de aprendizaje; peor aún, este fenómeno no es exclusivo del aprendizaje profundo y puede observarse incluso con clasificadores más simples. Además, dados algunos datos de entrada, parece ser relativamente fácil calcular una pequeña distorsión de esa entrada que será mal clasificada. Estas alarmantes observaciones tienen muchas implicaciones prácticas en un ámbito en el que las tecnologías de aprendizaje automático son ubicuas.

La investigación en curso en Inria se centra en parte en las redes neuronales convolucionales y sigue varias líneas. Es necesario comprender mejor los puntos débiles de las estrategias de aprendizaje profundo, analizar qué tipos de ataques son posibles y proponer mecanismos para proteger las redes neuronales convolucionales contra los ataques adversarios. Es importante comprender por qué las perturbaciones adversas se generalizan tan sorprendentemente bien. Otro enfoque consiste en monitorear las activaciones internas que fluyen entre las capas de una red profunda, con el fin de observar dónde se producen los elementos

adulterados, estén o no relacionados con la dimensionalidad de las representaciones vectoriales intermedias. Sin duda, merece la pena observar los efectos de estrategias defensivas como la destilación de estos flujos para ver por qué hace que el proceso general sea más robusto. Además, parece esencial conectar las perturbaciones adversarias con el subespacio que corresponde a las imágenes naturales, el subespacio donde se entrenan los clasificadores. Esto puede facilitar la identificación de imágenes no naturales que, de hecho, pertenecen a subespacios diferentes. Sin embargo, las imágenes naturales adversas existen y consiguen subvertir los sistemas de detección. Por lo tanto, considerar los subespacios no es la solución definitiva para hacer que los sistemas sean más robustos.

Por último, aunque la mayoría de las investigaciones se centran actualmente en las imágenes, otras modalidades como el texto adverso o el audio adverso, es probable que planteen toda una serie de nuevas dificultades que también hay que abordar. Es probable que los dominios ajenos a los multimedios se enfrenten a sensibilidades similares a los comportamientos adversos, ya que también dependen en gran medida del aprendizaje automático, por ejemplo, las aplicaciones para detectar intrusos a partir del análisis de la red o las aplicaciones que se ocupan de los rasgos biométricos.

CUESTIONES DE PRIVACIDAD Y APRENDIZAJE AUTOMÁTICO

El uso de técnicas de aprendizaje automático también plantea problemas de privacidad. Además del uso de técnicas de aprendizaje automático para inferir datos posiblemente sensibles, el aprendizaje automático también plantea la cuestión de si un atacante que tenga acceso a la red entrenada puede obtener información sobre los datos de entrenamiento. Se pueden distinguir diferentes escenarios según si un atacante tiene acceso a la red en 'caja blanca' (dando acceso a las partes internas de la red neuronal) o en 'caja negra', y si su objetivo es extraer los datos de entrenamiento o simplemente decidir si una entrada determinada formaba parte de los datos de entrenamiento. Estas propiedades de privacidad se han expresado en términos de privacidad diferencial, introducida en el ámbito de la anonimización de bases de datos (analizada en 5.2.2). En el peor de los casos, el atacante puede acceder a los propios datos de entrenamiento almacenados. Esto plantea las cuestiones adicionales de cómo transformar los datos antes de su almacenamiento, para descartar cualquier información privada que sea inservible para la tarea en cuestión, y cómo entrenar de forma distribuida en línea, para evitar el almacenamiento de todos los datos en un solo lugar, lo que aumenta el riesgo de una infiltración de la seguridad.

[Equipos Inria] Ciberseguridad y aprendizaje automático

Aunque el aprendizaje automático adverso se considera a menudo un problema de ciberseguridad, ya que puede utilizarse para atacar sistemas críticos que utilizan técnicas de aprendizaje automático, en realidad es bastante diferente de las técnicas tradicionales de ataque a la ciberseguridad y, por lo tanto, lo abordan principalmente los equipos que trabajan en el aprendizaje automático. Por el contrario, los problemas de privacidad que plantea el aprendizaje automático son típicamente de ciberseguridad.

↗ El equipo **COMETE** está interesado en los aspectos de privacidad del aprendizaje automático, en particular mediante el uso de la privacidad diferencial.

↗ El equipo **LACODAM**, que trabaja en el ámbito del aprendizaje automático, también está estudiando su aplicación a la ciberseguridad.

↗ El equipo **LINKMEDIA** investiga cuestiones de aprendizaje automático adverso que van más allá de la visión por computadora adversa y considera, por ejemplo, el audio adverso, el vídeo adverso, el reconocimiento automático del habla (ASR) adverso o el procesamiento del lenguaje natural (NLP) adverso. El aprendizaje automático adverso también se considera desde la perspectiva de cada modalidad, así como la consideración de entradas verdaderamente multimodales. El equipo también considera una amplia gama de objetivos de aplicación que van más allá de la clasificación e incluyen la recuperación tras un ataque.

↗ El equipo **MAGNET** trabaja en las consideraciones de privacidad en el aprendizaje automático, en el contexto del aprendizaje descentralizado en el que varios actores colaboran para mejorar el modelo sin filtrar datos personales, o en el minado de datos de los rastros de movilidad.

↗ El equipo **MULTISPEECH** y **MAGNET** trabajan en el reconocimiento del habla respetuoso con la privacidad. El objetivo es entrenar un sistema de reconocimiento del lenguaje en los datos del habla de los usuarios sin revelar información sobre la identidad, los rasgos (por ejemplo, el género, la edad o el origen étnico) o los estados (por ejemplo, la salud o el estado emocional) de los usuarios individuales. Para ello, los equipos investigan el aprendizaje adverso (que aquí se ve como una solución más que como un problema), el entrenamiento descentralizado y los marcos formales de privacidad como la privacidad diferencial.

↗ El equipo **ORPAILLEUR** está interesado en los aspectos de privacidad del aprendizaje automático.

↗ El equipo **PRIVATICS** y **TYREX** trabajan conjuntamente en el aprendizaje automático descentralizado o federado para preservar la privacidad, en particular en el contexto de las grandes bases de datos médicas.

↗ El equipo **SEQUEL** , que trabaja en el aprendizaje automático, también está estudiando aplicaciones a la ciberseguridad.

7

Ciberseguridad en Francia

[Resumen]
Francia es uno de los principales actores europeos en materia de ciberseguridad. Su personal académico está formado por casi mil personas, entre investigadores, profesores, becarios postdoctorales, estudiantes de doctorado e ingenieros de investigación. Casi el 25% de la actividad académica francesa en materia de ciberseguridad se lleva a cabo en Inria dentro de equipos conjuntos con otras instituciones académicas, principalmente el CNRS y las universidades, donde por término medio solo la mitad del personal es de Inria. Esto proporciona a Inria un efecto de palanca y la convierte en uno de los principales actores académicos franceses en materia de ciberseguridad. Las fuerzas de Inria en ciberseguridad representan alrededor del 7% de su actividad total. Un esfuerzo grande y visible, tanto en Inria como en Francia, se dedica a la criptografía y a los métodos formales aplicados a los protocolos criptográficos y a la privacidad. La seguridad del hardware también está bien representada en Francia, pero con muy pocos recursos de Inria. Los trabajos sobre seguridad de redes, seguridad de sistemas y, en general, sobre seguridad reactiva están, sin embargo, poco representados tanto en Inria como en Francia, tomando en cuenta los crecientes retos de ciberseguridad para las infraestructuras críticas y la llegada de la Internet de las cosas.
La educación en materia de ciberseguridad se está convirtiendo en un reto fundamental, a todos los niveles. Inria contribuye a la educación mediante el asesoramiento de estudiantes de doctorado, muchos de los cuales trabajarán posteriormente en organismos estatales o empresas privadas. La transferencia a través de la creación de start-ups existe, pero todavía no es un vector importante.
Las fuerzas académicas francesas en materia de ciberseguridad están bastante bien organizadas y coordinadas, en particular a través del grupo de trabajo pre-GDR y Allistene. Sin embargo, la interacción y la coordinación entre la industria y las instituciones académicas siguen siendo insuficientes. Aunque la ciberseguridad suele estar bien reconocida como una prioridad en Francia, en Europa y en todo el mundo, es importante que Francia también refuerce su apoyo financiero a mantener su posición de liderazgo y su capacidad para aprovechar las oportunidades económicas.

7.1 Fuerzas académicas en Inria y en Francia

Organización y medios de investigación

La investigación en Inria se organiza en pequeños equipos que comparten un proyecto de investigación común, que a menudo son equipos conjuntos con otras instituciones académicas como el CNRS, universidades, escuelas de ingeniería especializadas u otros institutos de investigación (INRA, INSERM, etc.). Por término medio, cerca de la mitad del personal de investigación de los equipos

de Inria procede de instituciones asociadas, lo que confiere a Inria un efecto de palanca. En lo sucesivo, la investigación en Inria significa siempre la investigación de los equipos de Inria, incluidos los socios. El tamaño medio de un equipo es de 18 personas –contando investigadores y profesores, becarios posdoctorales, estudiantes de doctorado e ingenieros de investigación–, pero con una gran variación, que va desde las 3-5 personas para los equipos más pequeños hasta las 45-50 personas para los más grandes, que son minoría.

La ciberseguridad ha sido una de las prioridades de investigación de Inria en los últimos quince años[127]. La ciberseguridad abarca ahora el 7% de la actividad de Inria, con unos 30 equipos trabajando en este campo[128], dos tercios de los cuales tienen la ciberseguridad como tema de investigación único o principal. En total, esto supone unos 200 puestos a tiempo completo. Esto representa una cuarta parte de las fuerzas académicas francesas en ciberseguridad. Las otras figuras académicas principales son los científicos del CNRS y los miembros de la docencia de las universidades y escuelas de ingeniería alojadas en las UMR del CNRS[129], pero fuera de los equipos de Inria, los miembros de la docencia del Institut Mines Telecom (IMT) y los científicos del CEA. Las fuerzas académicas francesas en ciberseguridad se han casi se han duplicado en la última década. Este crecimiento continúa, pero a un ritmo menor. Aunque hay algunas contrataciones nuevas, el crecimiento se debe sobre todo a los investigadores y profesores que se trasladan a la ciberseguridad desde otros campos de la informática (véase 1.1).

Ámbitos de investigación

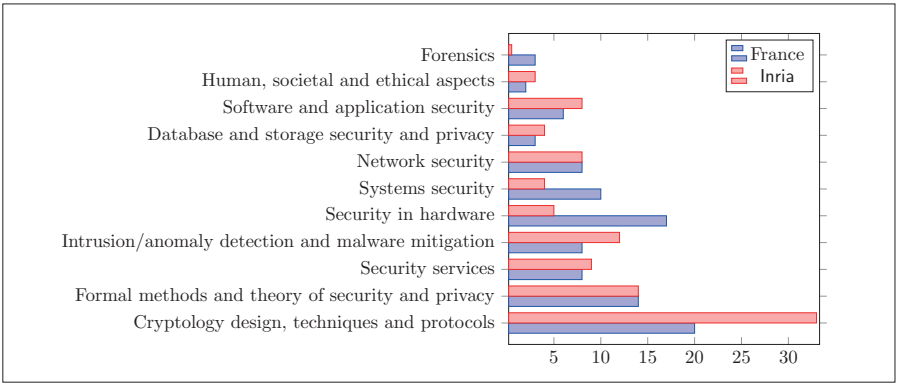


Figura 8.1: Principales temas de ciberseguridad por % de la fuerzas académicas francesas [130]

127. La seguridad se trató ampliamente en el Plan Estratégico de Inria para el periodo 2003-2007.
128. Todos los equipos de Inria que trabajan en ciberseguridad se enumeran en el Anexo A.
129. Unidades de investigación mixtas (organización).
130. Fuente: cartografía de las fuerzas académicas francesas en materia de ciberseguridad realizada por el grupo de trabajo sobre ciberseguridad de la Alianza Allistene, véase https://www.allistene.fr/files/2018/03/VF_cartographie_2017-06-13.pdf.

La figura 8.1 describe la proporción de actividades de investigación entre los principales ámbitos de la ciberseguridad en Inria, en rojo, y en el conjunto de entidades académicas francesas, en azul. La comparación es reveladora. La criptología es un punto fuerte de Inria, con un tercio de su plantilla. Esto es el resultado de la larga participación de Inria en la teoría de los números, el álgebra computacional, el criptoanálisis y la teoría de códigos. En efecto, Inria desempeña un papel clave a nivel mundial en el diseño de nuevas bases primitivas o protocolos criptográficos y en el criptoanálisis de primitivas criptográficas.

El siguiente ámbito de investigación más importante en Inria es el de los métodos formales aplicados a la seguridad y la privacidad. Como se explica en el apartado 1.1, esto se debe en gran medida a la transferencia de conocimientos procedentes de los métodos formales y al hecho de que Inria, y Francia en general, están muy bien posicionadas en el ámbito de los métodos formales. Por el contrario, la seguridad del hardware y de los sistemas está poco representada en Inria, pero este ámbito sigue estando bien cubierto en otras partes de Francia, en el CEA, el CNRS y el IMT. Hay que señalar que en Inria se investiga muy poco en el ámbito

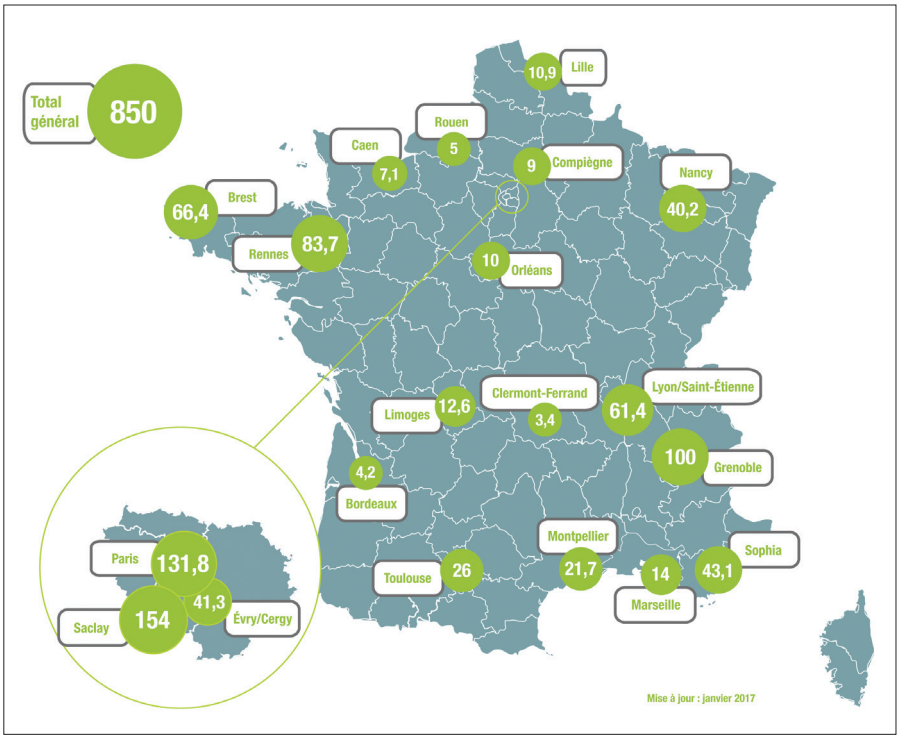


Figura 8.2: Desglose geográfico de las fuerzas académicas de ciberseguridad

forense. Sin embargo, en colaboración con tracIP[131], el equipo **RESIST** trabaja en el desarrollo de una plataforma forense dedicada a los sistemas de control industrial. Fuera de Inria, sí existe investigación en materia forense en Francia[132].

La figura 8.2 describe el desglose geográfico de las fuerzas de investigación académicas en Francia. La unidad de medida es el ETP, que equivale a una actividad de investigación a tiempo completo. Esto muestra que, además de París, las principales fuerzas en ciberseguridad están en la Bretaña (Brest y Rennes) y en la región de Ródano-Alpes (Lyon y Grenoble). A continuación, Nancy y la Costa Azul (Niza y Sophia Antípolis). Para Inria, las principales fuerzas están en la región de París (Île-de-France) y luego vienen Rennes y Nancy.

Fuerzas no académicas

Asimismo, hay importantes fuerzas de investigación no académicas. Es más, la investigación también se lleva a cabo en empresas privadas, aunque suele ser más aplicada. Por otra parte, existen organismos institucionales, como la DGA[133], que forma parte del Ministerio de Defensa francés, y la ANSSI[134], la agencia francesa de ciberseguridad, que forma parte de la SGDSN[135], dependiente de la secretaría del primer ministro, que cuentan con conocimientos de ciberseguridad de primer nivel, incluidas algunas actividades de investigación, y desempeñan un papel clave en la definición y dirección de las políticas francesas de ciberseguridad. Además, hay algunos laboratorios dependientes del Ministerio del Interior, como el CREOGN[136]. La investigación sobre cuestiones de privacidad y la normativa también se lleva a cabo en el laboratorio de innovación LINC[137] de la CNIL[138], la autoridad francesa de protección de datos. Los equipos de Inria mantienen regularmente colaboraciones científicas en algunos proyectos específicos con la mayoría de estas organizaciones.

Fomento de la actividad comunitaria

Adicionalmente, hay una serie de asociaciones que desempeñan un papel importante en fomentar la actividad en la comunidad de la ciberseguridad. El CNRS creó un pre-GDR[139] sobre ciberseguridad[140] en 2016, cuyo objetivo es animar a la comunidad académica francesa en materia de ciberseguridad, en particular mediante la organización de talleres o escuelas de verano.

131. https://www.tracip.fr
132. Es probable que se subestimen las actividades forenses en Francia, ya que las cifras no incluyen la investigación en humanidades.
133. Dirección General del Ejército.
134. Agencia Nacional de Seguridad de los Sistemas de Información.
135. Secretaría General de la Defensa y de la Seguridad Nacional.
136. Centro de Investigación de la Escuela de Oficiales de la Gendarmería Nacional.
137. Laboratorio de Innovación Numérica de la CNIL.
138. Comisión Nacional de Informática y Libertades
139. Un GDR (Groupement De Recherche) es una estructura dirigida por el CNRS para animar a la comunidad académica francesa de investigación en un área determinada.
140. http://gdr-securite.irisa.fr/index.html

La alianza Allistene[141] ha creado un grupo de trabajo sobre ciberseguridad[142] en el que Inria y los principales actores académicos tienen representantes y cuyo objetivo es intercambiar información, construir una visión común, realizar estudios de propósito general como la cartografía de fuerzas académicas descrita anteriormente, y coordinar acciones como la participación de los miembros de Allistene en el FIC[143].

Por último, Inria forma parte del grupo de trabajo dedicado a la investigación y la innovación del CoFIS (Comité de la Industria de la Seguridad), cuya función principal es fomentar la industria francesa de la seguridad proponiendo acciones específicas para aumentar tanto la competitividad como la seguridad a nivel nacional y europeo.

Inria también es miembro de la asociación profesional ACN[144], cuya función es agrupar y representar a los principales actores industriales de la ciberseguridad. HEXATRUST es otra importante asociación compuesta por 29 PYMES en ciberseguridad, donde sin embargo no están representados Inria ni la investigación académica.

7.2 Educación

La educación es un tema importante, ya que existe una enorme laguna de conocimientos en materia de ciberseguridad a todos los niveles, pero los conocimientos de alto nivel siguen siendo el quid de la cuestión. Esta situación se conoce desde hace unos años, por lo que algunas escuelas de ingeniería y universidades han creado nuevos programas de ciberseguridad. Aunque Inria no es una universidad, la educación y la transferencia de conocimientos siguen siendo una de sus misiones importantes, e Inria contribuye a la educación de varias maneras.

Muchos investigadores de Inria imparten cursos avanzados de ciberseguridad en escuelas de ingeniería y universidades, especialmente en cursos de nivel de máster, y en escuelas de verano especializadas. Inria también ha creado un par de MOOC sobre ciberseguridad. Y lo que es más importante, la mayoría de los investigadores están asesorando a estudiantes de doctorado. Los que no continúan en el mundo académico serán contratados por la industria u otras instituciones; esta es una de las formas más eficientes de transferir conocimientos. El número de estudiantes de doctorado es más o menos el mismo que el número total de investigadores y profesores, una proporción que es aproximadamente la misma para la ciberseguridad y otros ámbitos de investigación. Dado que la industria carece de expertos en ciberseguridad, es importante aumentar el número de

141. http://www.allistene.fr/
142. https://www.allistene.fr/organisation-allistene/groupes/groupe-cybersecurite/
143. Foro Internacional de la Ciberseguridad.
144. Alliance pour la Confiance Numérique.

estudiantes de ciberseguridad a todos los niveles, incluidos los de doctorado. Inria podría aumentar su número de doctores, pero sólo si hay un aumento correspondiente en la financiación de doctorados en ciberseguridad y en buenos candidatos a doctorado. Por tanto, el principal reto sigue siendo atraer a más estudiantes jóvenes a la informática en general y a ciberseguridad en particular (véase el apartado 8.2.4).

7.3 El impacto de Inria en la ciberseguridad

Francia es uno de los países líderes mundiales en ciberseguridad. Esta fortaleza se debe a sus fuerzas académicas, incluidas las de Inria. Además de los sólidos resultados de la investigación y de mantener el más alto nivel de experiencia en la mayoría de los subdominios de la ciberseguridad, la investigación de Inria es también una contribución clave para la comunidad.

Francia e Inria llevan mucho tiempo participando en retos de criptoanálisis, ostentando varios récords de factorización. Este esfuerzo es necesario para comprobar continuamente el estado del arte del criptoanálisis, tanto en términos de algoritmos como de potencia de cálculo, y para recomendar, en consecuencia, ajustes de las claves criptográficas o un cambio de las bases primitivas criptográficas. Esta es una contribución clave para la comunidad.

Los investigadores también suelen formar parte de comités de normalización, como el IETF. Esto suele corresponder a un esfuerzo a largo plazo que consume mucho tiempo, pero es importante para mejorar la calidad de las normas.

Los equipos emergentes de Inria

En comparación con el impacto de su investigación en ciberseguridad, sólo unas pocas start-ups en ciberseguridad están surgiendo directamente de los equipos de Inria.

CRYPTOSENSE

Fundada en 2013 y con sede en el centro de París, Cryptosense[145] es una spin-off académica derivada de Inria y la Universidad Ca'Foscari de Venecia. Desarrollan software basado en la investigación académica llevada a cabo por Graham Steel y sus colegas del equipo Prosecco y los antiguos equipos Secsi.

Su principal producto, Cryptosense Analyzer, descubre fallos de seguridad en los sistemas criptográficos mediante diversas técnicas. El hardware criptográfico se trata como una caja negra y se prueba mediante técnicas de fuzzing. Los resultados de fuzzing se utilizan para inferir un modelo lógico del dispositivo que luego se analiza mediante la verificación de modelos. El software criptográfico se

145. https://cryptosense.com/

comprueba utilizando llamadas a las bibliotecas criptográficas. A partir de ellas se infiere un modelo que se analiza con respecto a una base de datos de reglas de uso criptográfico. Las propias bibliotecas se someten a pruebas de un conjunto de ataques criptográficos.

El software de Cryptosense es utilizado por varios bancos internacionales, proveedores de pagos, empresas tecnológicas, organismos gubernamentales y fabricantes de hardware, en Europa y Norteamérica.

CYBER-DETECT

La start-up Cyber-Detect[146] se creó a partir de la investigación sobre el malware realizada en el antiguo equipo Carte y continuada en el equipo Carbone LORIA. Desde hace varios años, desarrollaron una solución, denominada análisis morfológico, para analizar los códigos binarios y detectar programas malignos. El análisis morfológico es un método que consiste en abstraer el gráfico de flujo de control de un código binario y construir automáticamente firmas a partir de esta abstracción. La recombinación de firmas permite identificar las funcionalidades maliciosas. El prototipo resultante es ahora comercializado por Cyber-Detect como una solución de ayuda al análisis inverso y al análisis forense.

LYBERO.NET

La start-up Lybero.net[147], surgida del centro de investigación Inria Nancy – Grand Est, propone dos servicios comerciales. El primero es una custodia digital descentralizada basada en el quórum que permite la recuperación de una clave olvidada (o cualquier contenido digital) si se ha reunido un quórum preestablecido de administradores.

El segundo servicio es CryptnDrive, un controlador criptográfico seguro que permite un intercambio seguro y sencillo de archivos dentro y fuera de una organización. A través de un navegador web normal, los usuarios pueden intercambiar de forma segura documentos importantes sin tener que enviar ninguna contraseña al destino remoto, garantizándose la confidencialidad mediante un cifrado de extremo a extremo. Además, este servicio aprovecha la custodia basada en el quórum para recuperar las credenciales perdidas.

MALIZEN

Aunque se utilizan muchas herramientas para asegurar nuestros sistemas de información de forma proactiva, vemos regularmente que los atacantes encuentran formas de eludirlas o de explotar sus debilidades. Las soluciones de monitorización permiten detectar, caracterizar y responder a estas intrusiones en la mayoría de los casos. Sin embargo, a veces es necesario gestionar las excepciones o verificar

146.https://www.cyber-detect.com/
147. https://lybero.net/

el buen funcionamiento de estos sistemas. Frente a las ingentes cantidades de datos recogidos por la vigilancia, los expertos en seguridad suelen estar mal equipados y tienen dificultades para responder a los incidentes de seguridad.

Impulsada por los prometedores resultados de los proyectos de investigación en colaboración entre Inria y la DGA-MI, Malizen[148] es una start-up cuyo objetivo es equipar a los expertos en ciberseguridad con paquetes de hardware y software para ayudarles a responder mejor a los incidentes de seguridad.

Mediante el uso de interfaces gráficas de usuario y visualización de datos, especialmente diseñadas para la ciberseguridad, los expertos se reintegran en el proceso de análisis. Los responsables de analizar las intrusiones pueden explorar sus datos de seguridad de forma más intuitiva y comprender mejor las situaciones críticas.

Malizen cuenta con el apoyo de Inria y la investigación del equipo **CIDRE**.

7.4 Laboratorios de alta seguridad (LHS)

Inria tiene también dos Laboratorios de Alta Seguridad (LHS) situados en los centros de investigación de Nancy y Rennes que son compartidos con los socios locales de Inria, el CNRS y las universidades, y que también cuentan con el apoyo de las regiones de Gran Este y de Bretaña, y de la DGA en Rennes.

Ambos LHS se encuentran en salas seguras. El LHS de Nancy tiene un telescopio de red que captura el código del malware y los registros de los ataques y permite experimentar con sondas de Internet. Dispone de una red cerrada para experimentos sensibles, como el análisis de código de malware, y de una sala aislada, no conectada a Internet, donde se puede procesar información altamente sensible y realizar experimentos confidenciales de hardware y software. El Rennes-LHS alberga tres plataformas: una para la experimentación con la infección y reparación de malware y ransomware, que contiene una base de datos de virus informáticos y ransomware; una para la observación y el análisis electromagnético con el fin de experimentar con ataques de canal lateral; y la última para la inducción de fallos en el hardware electromagnético.

148. https://malizen.com/

8

Conclusiones y recomendaciones

Mientras que hace dos décadas la ciberseguridad no era un problema para el público en general, hoy en día es algo que concierne a todos: Estados, industrias y ciudadanos. La ciberseguridad es un tema candente con muchas consideraciones de seguridad económica, social, política o geopolítica en juego, y muy probablemente seguirá siéndolo en las próximas décadas. Concluimos este libro blanco con la lista de retos de investigación que hemos identificado y algunas recomendaciones generales relacionadas con la organización de la investigación en ciberseguridad y la interacción entre los investigadores en ciberseguridad y la sociedad.

8.1 Retos de la investigación

Aunque los retos identificados a lo largo de este libro blanco no son los únicos temas importantes, los consideramos de especial importancia y recomendamos que se traten de forma prioritaria. Por ello, y por razones de utilidad, los citamos a continuación. Cada reto puede ponerse en contexto en la sección en la que se expuso por primera vez.

8.1.1 Ataques de software dirigidos al hardware *(véase 2.1)*

Los ataques contra los sistemas de información no suelen afectar a la capa de hardware, sino que explotan una vulnerabilidad del software. Sin embargo, ataques recientes, como Rowhammer, Spectre o Meltdown, han demostrado que los ataques implementados en software pueden explotar las optimizaciones de rendimiento del hardware. Este nuevo tipo de ataque es especialmente peligroso, ya que hace posible las agresiones al hardware a distancia, a diferencia de los clásicos ataques de canal lateral. Todavía no está del todo claro cómo pueden las actuales intrusiones en fase de 'prueba de concepto' expandirse a nivel industrial, pero abren el camino a una nueva clase de ataques graves. Por lo tanto, es necesario comprender mejor cómo podrían desplegarse estas agresiones, proponer una tipología clara de este nuevo tipo de ataque y plantear contramedidas, tanto a nivel de hardware como de software. Esta tarea requiere conocimientos especializados a nivel de hardware, firmware y sistema operativo. Las contramedidas también pueden ser difíciles de diseñar, ya que pueden requerir la revisión de optimizaciones cruciales utilizadas durante años, como la ejecución especulativa.

8.1.2 Seguridad y usabilidad *(véase 2.3.2)*

Muy a menudo, cuando los usuarios solicitan un servicio, están dispuestos a sacrificar la seguridad, y a saltarse un mecanismo de seguridad molesto, si eso les impide utilizar el servicio. Para evitar este problema, la seguridad debe ser lo más transparente posible. Aunque la transparencia total no siempre es posible, los servicios de seguridad deben ser lo más sencillos posible de utilizar. Hay que trabajar para proponer interfaces y mecanismos de seguridad que se adapten a los

usuarios no expertos, que garanticen que el usuario conoce bien las consecuencias de sus acciones y que eviten que los usuarios cometan errores que comprometan la seguridad. El diseño de estos mecanismos de seguridad aplicables exige una investigación interdisciplinar que suele incluir a expertos en ciencias cognitivas.

8.1.3 Criptografía postcuántica *(véase 3.1.3)*

Se cree que la construcción de una computadora cuántica universal (y no para un propósito específico) será factible en las próximas décadas. Por tanto, es importante pensar ahora en una criptografía resistente a una computadora cuántica, ya que cierta información que se cifra hoy puede seguir siendo sensible dentro de, por ejemplo, 50 años. La mayor parte de la criptografía asimétrica utilizada hoy en día se basa en la solidez de la factorización o en el procesamiento de logaritmos discretos, problemas que se sabe que una computadora cuántica puede resolver con facilidad. Por lo tanto, es necesario buscar alternativas: las primitivas basadas en retículos, en códigos y en multivariantes son las candidatas más destacadas. Es urgente realizar un análisis de seguridad en profundidad de estos nuevos planteamientos.

8.1.4 Computación sobre datos encriptados *(véase 3.2.2)*

La necesidad de computar sobre datos encriptados ha surgido, en particular, con la aparición de la nube y la computación externalizada. En criptografía, este problema puede resolverse mediante un cifrado homomórfico o funcional. En 2009, Gentry demostró en su innovador artículo que era posible construir un esquema de cifrado totalmente homomórfico (FHE). Sin embargo, esta construcción seguía siendo teórica y resultaba completamente inviable debido a su escaso rendimiento. Desde entonces, han habido avances importantes en los esquemas FHE, logrando una velocidad aún muy baja de 50 puertas lógicas por segundo, aproximadamente. Los avances significativos tendrán aplicaciones extremadamente útiles para la computación en la nube que preserva la privacidad, donde cualquier progreso técnico puede ser rápidamente aprovechado económicamente.

8.1.5 Protocolos criptográficos de extremo a extremo formalmente verificados *(véase 3.3.4)*

Dado que la seguridad de los protocolos criptográficos es extremadamente difícil de garantizar (las verificaciones de lápiz y papel suelen contener errores), el uso de métodos rigurosos y formales aparece cada vez más como la única manera de alcanzar el nivel de seguridad esperado para esta clase de sistemas. Por lo tanto, el ámbito de las pruebas de seguridad asistidas por computadora es un tema cada vez más importante y debe incluir todos los aspectos, desde la especificación hasta la implementación. Los trabajos recientes, en particular en torno a TLS 1.3, han demostrado que esto ya es posible.

Sin embargo, estas pruebas siguen requiriendo un código cuidadosamente elaborado y un nivel muy alto de conocimientos. Aprovechar las técnicas de demostración para hacerlas aplicables a un código más general y utilizables por un público más amplio es ahora el principal reto. Los distintos protocolos suelen garantizar propiedades de seguridad diferentes, pero las herramientas existentes para verificar ciertas propiedades, como el anonimato, aún no tienen la misma madurez que las herramientas para verificar las propiedades de autenticación. Otro reto es considerar modelos de adversarios más fuertes, por ejemplo, un adversario que pueda controlar parte de la computadora a través de un programa maligno.

8.1.6 Detección de intrusos en redes encriptadas *(véase 4.4.1)*

Hoy en día, la detección de intrusos se realiza esencialmente a nivel de red. Si, como se espera en un futuro próximo, el tráfico se encriptara más sistemáticamente, lo que por supuesto sería una buena práctica para la seguridad y la privacidad, el análisis de los paquetes de red se volvería de facto inoperante, excepto para el análisis de la cabecera de los mensajes. Por lo tanto, se hace importante estudiar y diseñar nuevos mecanismos de supervisión de los sistemas de información y de producción de alertas, a nivel de aplicación, de middleware, de sistema operativo, e incluso de firmware o de hardware.

8.1.7 Comprender la privacidad y obtener herramientas prácticas *(véase 5.1.6)*

Comprender los principios y la normativa sobre privacidad es la base de cualquier actividad en este ámbito. Aunque no se trata de un ámbito nuevo (por ejemplo, la "Loi Informatique et Libertés", o Ley Informática de Protección de Datos, se adoptó en 1978), este ámbito ha experimentado recientemente importantes avances con el nuevo reglamento europeo RGPD y, al mismo tiempo, nuevas oportunidades de recopilar datos personales. En consecuencia, la comprensión de los conceptos y de la normativa es una primera necesidad. Otra de ellas es poder obtener herramientas prácticas: aunque el RGPD promueve varios conceptos y objetivos, proporciona poca orientación sobre la aplicación eficaz de estas nuevas disposiciones reglamentarias.

En particular, el RGPD introdujo el derecho a la portabilidad de los datos, en virtud del cual un usuario puede recuperar sus datos en un formato legible por humanos y por máquinas. Este derecho abre nuevos campos de investigación en torno a la gestión y el control individualizados de los datos personales. El objetivo es capacitar a los ciudadanos para que aprovechen sus datos personales para su propio bien, lo que exige plataformas personales en la nube seguras, extensibles y soberanas, tres objetivos contrapuestos que abren nuevos retos de investigación (véase, por ejemplo, el apartado 5.2.4).

8.1.8 Datos abiertos y anonimización *(véase 5.2.3)*

Las iniciativas de datos abiertos pueden implicar a veces la publicación de

bases de datos que contienen información personal sensible. Para garantizar la privacidad de los individuos, los datos deben ser anonimizados. La anonimización robusta, que resiste eficazmente los ataques de supresión del anonimato, es un tema de investigación activo y candente. Aunque la privacidad diferencial se ha convertido en una herramienta científica clave para lograr garantías de anonimización comprobables, siguen existiendo retos en su aplicación, por ejemplo, para mejorar la relación privacidad/utilidad.

8.1.9 Hacia un mundo conectado inteligente que preserve la privacidad *(véase 5.3.6)*

Nuestro mundo conectado experimenta un crecimiento sin precedentes en cuanto a la recogida de datos personales, con prácticas cada vez más intrusivas para la intimidad del ciudadano. Navegar por la web, utilizar smartphones y otros dispositivos inteligentes, conducir un auto conectado –y pronto autónomo– son actividades que generan filtraciones de datos personales. La falta de transparencia (muchos servicios y dispositivos se comportan como cajas negras) y la falta de control del usuario (cómo expresar su consentimiento u oposición cuando no hay ni información ni interfaz de usuario) son problemas importantes.

La identificación de estos comportamientos ocultos se ve dificultada por el número y la complejidad de las tecnologías subyacentes específicas de cada ámbito. Por ejemplo, la identificación de las prácticas de seguimiento en una página web requiere análisis avanzados de la ejecución de JavaScript, mientras que la supervisión de las aplicaciones de los teléfonos inteligentes necesita marcos específicos, y la supervisión de determinadas tecnologías de comunicación inalámbrica sigue sin resolverse en su mayor parte. El análisis de estos flujos de datos es necesario para evaluar posibles pérdidas de privacidad, por ejemplo, en un hogar inteligente.

Estas desafiantes y heterogéneas actividades de investigación son esenciales para aportar transparencia, poner de relieve las buenas y malas prácticas y permitir a los reguladores hacer cumplir las leyes de protección de datos. Como tal, esta investigación ayuda directamente a dar forma a nuestro futuro mundo conectado inteligente.

8.1.10 Asegurar la Internet de las Cosas (IoT) *(véase 6.2.1)*

La seguridad en la IoT es un reto importante: los ataques son todavía relativamente fáciles (muchos dispositivos no se han diseñado teniendo en cuenta la seguridad), invasivos (por ejemplo, omnipresentes en nuestras vidas) y con un impacto potencialmente importante debido al factor de multiplicación propiciado por el gran número de dispositivos disponibles y a las interacciones directas que algunos de ellos tienen en el mundo físico (por ejemplo, los autos conectados). Las líneas de investigación son múltiples, como por ejemplo la capacidad de actualizar de forma segura el software de los dispositivos integrados, el diseño

de primitivas criptográficas ligeras adaptadas a recursos limitados, el análisis de la seguridad de las nuevas tecnologías inalámbricas de baja potencia y de área amplia, la detección y mitigación de intrusiones o dispositivos que se comportan erróneamente, y la necesidad de marcos, protocolos y sistemas operativos seguros desde el diseño.

8.1.11 Sistemas industriales seguros *(véase 6.2.2)*

Los sistemas industriales se basan cada vez más en mecanismos de software que pueden ser atacados. Por ello, su seguridad se ha convertido en un problema importante, especialmente porque las consecuencias de un ataque contra estos sistemas pueden ser dramáticas. Aunque los enfoques de seguridad tradicionales parecen aplicables al caso de los sistemas industriales, sus especificidades exigen una revisión de los mecanismos de seguridad tradicionales para adaptarlos a este nuevo contexto. En particular, los protocolos de comunicación utilizados en este contexto no pueden modificarse de la noche a la mañana. Debe haber una transición durante la cual las comunicaciones heredadas deben integrarse en protocolos seguros. Además, suele ser necesario controlar el sistema en tiempo real. Por tanto, la seguridad también debe ser aplicable en tiempo real. Por último, a menudo es imposible modificar los dispositivos industriales. Por lo tanto, no se pueden utilizar mecanismos de seguridad preventiva. Por ello, es obligatorio utilizar la seguridad reactiva y, en consecuencia, es muy importante estudiar cómo se pueden desplegar mecanismos de detección de ataques eficaces en este contexto.

8.2 Recomendaciones generales

A continuación ofrecemos algunas recomendaciones generales relacionadas con la interacción entre los investigadores en ciberseguridad y la sociedad, la organización de la investigación en ciberseguridad y, para concluir, la importancia de la ciber-resiliencia.

8.2.1 La sociedad debería beneficiarse más de la experiencia científica académica

Para la mayoría de los temas de ciberseguridad, hay académicos con conocimientos técnicos muy específicos que pueden proporcionar, y a veces lo hacen, un asesoramiento científico útil. Sin embargo, los científicos suelen estar poco representados en los comités consultivos nacionales o industriales, en comparación con los miembros de la industria. Además, algunas posiciones o decisiones adoptadas, a diferentes niveles, muestran una falta de asesoramiento científico.

8.2.2 Transferencia de conocimientos entre la ciberseguridad y otros ámbitos

La necesidad de conocimientos en materia de ciberseguridad es sorprendente en la mayoría de sus ámbitos de aplicación, por ejemplo, los sistemas industriales, los sistemas médicos, la robótica y, lo que es quizá más importante, la IoT (véase la lista completa en el apartado 6). Desgraciadamente, la ciberseguridad aún no ha sido suficientemente identificada como una prioridad en estos ámbitos y, por lo tanto, no se tiene en cuenta en las primeras fases de diseño de las aplicaciones, lo que impide respetar el principio de seguridad por definición. Por el contrario, la investigación en ciberseguridad a veces necesita más experiencia en métodos formales o inteligencia artificial, por ejemplo. Para invertir esta tendencia, es necesario que haya una transferencia de conocimientos entre los equipos de investigación en ciberseguridad y estos otros ámbitos.

8.2.3 Promover la seguridad también como ciencia experimental

La seguridad de los sistemas y la seguridad de las redes parecen adolecer de falta de prestigio en el mundo académico, al menos en Francia. Esto puede ser una cuestión cultural, en parte debido a que la investigación en estos ámbitos es más experimental y tecnológica. Otra preocupación es la falta de conjuntos de datos del mundo real. Por ejemplo, la detección de intrusos y la correlación de alertas adolecen de una falta de datos reales en los que sea posible probar y comparar los nuevos mecanismos propuestos por los investigadores. Ayudar a los investigadores a acceder o generar esos datos parece crucial.

8.2.4 Educación

La educación es esencial para la seguridad (véanse los apartados 1.3 y 2.3.3) y deben realizarse importantes esfuerzos de divulgación para todos los públicos: profesores, personal docente, investigadores, agentes industriales y especialistas, hasta los ciudadanos de a pie, incluidos los niños pequeños. En lo que respecta a la concienciación, es necesaria la mediación científica en materia de ciberseguridad, dirigida a la sociedad en general, y con especial atención a los académicos y profesores de los centros de enseñanza primaria y secundaria. En el ámbito profesional, la formación de doctorado en áreas de ciberseguridad es, por supuesto, una herramienta educativa lógica. De todas maneras, parece recomendable que los másteres profesionales especializados o los programas de último año de las escuelas de ingeniería ya aportaran a los estudiantes una visión clara del posible futuro de la seguridad de la información (más allá del estado del arte de la industria). En este contexto, los MOOCs podrían ser un vehículo de difusión y educación eficiente, así como una importante herramienta de comunicación, ya que potencialmente llegan a varios miles de usuarios por cada sesión de MOOC y requieren una participación activa por su parte.

8.2.5 Ciber-resiliencia por definición

Muchos informes recientes sobre el análisis de las amenazas a la ciberseguridad afirman que los Estados tendrán que hacer frente a ciberataques masivos. ¿Estamos bien preparados para resistir un "cibertornado"? La resiliencia de las infraestructuras críticas y, en particular, de los operadores de vital importancia (OVI) es un problema real. En realidad, gran parte de este reto es organizativo y técnico, más que académico, y la ANSSI ya ha realizado muchos esfuerzos en aumentar nuestra resiliencia. Sin embargo, la investigación puede ayudar: el principio de seguridad por definición debería aplicarse también a la ciber-resiliencia. Esto significa que no debe ser planteada a posteriori, sino que debe tenerse en cuenta desde la fase inicial del diseño de los sistemas digitales, redes e infraestructuras digitales.

Además, muchas subdivisiones de la ciberseguridad también están indirectamente implicadas en este desafío: la seguridad reactiva o la detección de malware son de suma importancia; la seguridad preventiva para garantizar los niveles de protección más avanzados también es relevante; los métodos formales, incluidas las pruebas de protocolos, sus implementaciones, así como las partes críticas de los sistemas operativos, pueden tener un impacto significativo a largo plazo en la resiliencia de los sistemas de comunicación.

Observaciones finales

En los últimos años se ha realizado un importante esfuerzo nacional para reforzar la seguridad del Estado y de los sistemas de información de los operadores vitales. En particular, la ANSSI (Agence Nationale pour la Sécurité des Systèmes d'Information) y la DGA (Délégation Générale de l'Armement) –por citar solo dos ejemplos– han contratado a muchos expertos para adaptar sus recursos humanos a las necesidades a cubrir. Sin embargo, aunque cada vez se reconoce más la importancia de la investigación y la formación en materia de ciberseguridad, se han hecho muy pocos esfuerzos para contratar a más profesores e investigadores. En el ámbito de la ciber-resiliencia, los recursos humanos son esenciales: por lo tanto, pedimos que continúen los esfuerzos emprendidos por algunas instituciones nacionales, y que estos esfuerzos se extiendan a las instituciones de investigación, universidades y escuelas.

Del mismo modo que los países están realizando importantes esfuerzos para mejorar sus capacidades en materia de ciberseguridad, es fundamental que Francia mantenga o incluso refuerce sus fuerzas académicas y todo su ecosistema de innovación en este campo.

Anexo A

Lista de equipos

LEYENDA: Los equipos cuyo nombre está todo en mayúsculas desarrolla su actividad principal en relación con la ciberseguridad. Otros equipos para los que solo la letra inicial del nombre está en mayúsculas y con tipografía normal desarrollan alguna actividad significativa en el área de la ciberseguridad, mientras que los equipos cuyo nombre comienza en mayúscula y con letra en cursiva solo realizan una contribución marginal a la ciberseguridad. Los sitios web de los equipos y los informes anuales de 2017 pueden encontrarse en las URL https://www.inria.fr/en/research/research-teams y https://raweb.inria.fr/rapportsactivite/RA2017/.

Domaine de Voluceau, Rocquencourt BP 105
78153 Le Chesnay Cedex, France
Tél. : +33 (0)1 39 63 55 11
www.inria.fr