Seguridad y Privacidad de Datos en Sistemas de Ciencia de Datos en Salud

Melissa Brigitthe Hinojosa-Cabello[0000−0002−0404−0398], Miguel Morales-Sandoval[0000−0003−1702−8467], Diana Elizabeth Carrizales-Espinoza[0000−0002−3925−031X], y José Luis González-Compeán[0000−0002−2160−4407]

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Ciudad Victoria, Tamaulipas, 87130, México
{melissa.hinojosa,miguel.morales,diana.carrizales, joseluis.gonzalez}@cinvestav.mx

Resumen El término *Big Data* se refiere a la producción de datos a gran velocidad, de una gran variedad y con un alto volumen. En el ámbito de la salud, *Big Data* se refiere a los procesos involucrados en la gestión, almacenamiento, tratamiento y uso de datos médicos que pueden provenir de distintas fuentes. En este contexto, paradigmas como el Internet de las Cosas Médicas (IoMT) o el cómputo en la nube han acelerado la producción masiva de datos en el área de la salud. Los datos originados en nodos IoMT (e.g., monitoreo de frecuencia cardiaca, niveles de glucosa, etc.) se almacenan en la nube y después son consumidos o accedidos desde la misma por los usuarios finales (pacientes, profesionales de la salud, personal médico, entre otros) a partir de diversas aplicaciones. Dada la naturaleza sensible de los datos médicos, que incluyen información tanto personal como médica de los pacientes, resulta necesario preservar el derecho a la privacidad. Si hay datos sensibles, éstos deben estar protegidos en todo momento, durante el ciclo de vida de los datos, ante cualquier divulgación o modificación no autorizada. Es decir, los propietarios de los datos esperan que éstos únicamente estén disponibles y puedan ser accedidos por usuarios autorizados, sin que el proveedor del servicio u otras entidades no autorizadas sean capaces de obtener y procesar dichos datos. En este sentido, los requerimientos de seguridad que se deben cubrir son confidencialidad y control de acceso, principalmente.
En este capítulo presentamos una descripción y detalles de construcción del concepto de *sobres digitales con capacidades de búsqueda*, los cuales son objetos criptográficos que permiten garantizar la privacidad de datos sensibles, como los de salud. Así, los datos únicamente serán accedidos por entidades autorizadas, descritas por un conjunto de atributos que los caracterizan e identifican. De igual forma, el almacenamiento y recuperación segura de datos médicos es indispensable en el desarrollo de sistemas de ciencia de datos. Por ello, al final de este capítulo se describe un caso de uso de los sobres digitales en este tipo de aplicaciones.

Palabras Clave: Confidencialidad · Control de Acceso · Sobres Digitales · eSalud · Big Data.

1. Introducción

La seguridad informática se refiere a todos aquellos mecanismos y recursos utilizados para prevenir accesos no autorizados a los sistemas de información, que incluyen recursos o infraestructura de cómputo, sistemas y datos. De entre éstos, la seguridad de datos es la última línea de defensa, ya que cuando un atacante logra romper la seguridad de la red y del dispositivo, éste tiene la posibilidad de acceder a los datos y comprometer su confidencialidad. En este capítulo, a menos que se indique lo contrario, nos enfocamos únicamente en la seguridad de los datos. Desde esta perspectiva, la seguridad de datos la abordamos desde dos requerimientos principales: confidencialidad de datos y control de acceso hacia éstos.

Definición 1. *Confidencialidad* [6], [7]: La confidencialidad garantiza la privacidad de datos sensibles al impedir su divulgación mediante la restricción del acceso a éstos a personas, recursos o procesos no autorizados, permitiendo que únicamente aquellos con autorización legítima puedan acceder a los datos, consumirlos o procesarlos. Éste es el requerimiento más antiguo y también el más demandado cuando se habla de seguridad de datos.

Definición 2. *Control de acceso* [7]: El objetivo del control de acceso lógico es la protección de cualquier tipo de recurso (datos, aplicaciones, servicios, entre otros) de operaciones inadecuadas llevadas a cabo por usuarios malintencionados. Éste consiste en la definición de una serie de restricciones, normalmente basadas en políticas, que describen quién puede acceder a los recursos y las operaciones permitidas sobre éstos, e impiden el acceso no autorizado mediante soluciones tecnológicas. El control de acceso involucra herramientas y protocolos para gestionar el acceso a sistemas y recursos mediante la identificación, autenticación y autorización de los usuarios.

La confidencialidad de datos puede alcanzarse mediante el cifrado de los mismos. Cifrar significa, a grandes rasgos, una transformación de los datos (D), de un formato legible a uno ilegible (CT), mediante un procedimiento (P) bien definido y conocido. Para realizar dicha transformación se usa una llave criptográfica k_c que, en términos simples, corresponde a una secuencia de bits de longitud n con suficiente aleatoriedad. Este proceso de cifrado se representa por la Ecuación 1. Una vez cifrados, los datos D ya no son accesibles por nadie, salvo por aquellos que posean una llave para descrifrar k_d, y mediante un proceso inverso al cifrado (i.e., descifrado, P^{-1}) puedan transformar nuevamente los datos cifrados CT en D, como se muestra en la Ecuación 2.

$$CT = P(D, k_c) \tag{1}$$

$$D = P^{-1}(CT, k_d) \tag{2}$$

En este sentido, la premisa del cifrado es que, para cualquier entidad que desconozca k_d resulta prácticamente imposible obtener los datos legibles a partir del

texto transformado CT tras ejecutar P^{-1}, incluso siendo éste un procedimiento bien conocido. Todos los cifradores, tanto antiguos como modernos, basan su funcionamiento en los preceptos previamente descritos y, en principio, garantizan el servicio de confidencialidad. Una vez cifrados, los datos pueden almacenarse en un medio inseguro (e.g., en una unidad de disco), transmitirse también por un medio inseguro (como internet), o enviarse a la nube. Sin embargo, es necesario que el propietario de datos imponga y maneje las restricciones de acceso a través de un control sobre las llaves de descifrado. Ante un escenario donde existen grandes colecciones de datos que deben cifrarse para garantizar su confidencialidad, como en el caso de *big data* en salud, se deben resolver al menos tres problemáticas principales: eficiencia, compartición y recuperación.

- La *eficiencia*: El cifrado, finalmente, es un requerimiento no funcional que convella una sobrecarga, tanto en procesamiento como en almacenamiento, para ejecutar los procedimientos P y P^{-1}. El nivel de seguridad está correlacionado con la longitud de la llave k_c, por lo que entre más grande k_c, mayor es el nivel de seguridad, pero también más lentos los procedimientos P y P^{-1}.
- La *compartición*: Los datos generalmente no son consultados solo por el propietario de los mismos. En el caso de eSalud, los datos deben ser accedidos por distintos actores, como los mismos pacientes, médicos, enfermeros, especialistas y, en general, profesionales de la salud. Ante un creciente volumen de datos se hace evidente la necesidad de contar con mecanismos efectivos y eficientes de control de acceso hacia dichos datos, para una compartición no solamente segura, sino también eficiente.
- La *recuperación*: Dado un gran volumen de datos, la recuperación de información es necesaria para localizar rápidamente datos de interés y recuperarlos para su acceso y consumo. Pero, ¿qué pasa si los datos están cifrados y, por lo tanto, se encuentran en formato ilegible? ¿Cómo un motor de búsqueda puede localizar y recuperar datos ilegibles de interés? Éste es uno de los desafíos más relevantes en el contexto del cifrado de grandes colecciones de datos. Afortunadamente, existen mecanismos que permiten abordar este problema, en lo que se conoce como *Searchable Encryption (SE)* [21]. Bajo este enfoque, los usuarios de los datos pueden hacer búsquedas cifradas, esto es, enviar al proveedor del servicio de almacenamiento de los datos cifrados un *'token'* que indica palabras clave en formato cifrado, de tal forma que éste pueda usarlo para buscar en los datos cifrados y localizar aquellos que empaten con los criterios de búsqueda. Al estar cifrados el *token* y los datos, el servidor realiza algo parecido a una búsqueda a ciegas pero efectiva, recuperando los datos de interés sin aprender acerca de los criterios de búsqueda o de los datos localizados. Sin embargo, no cualquier esquema SE podría ser adecuado para un entorno particular.

En este capítulo presentamos el concepto, diseño, implementación y evaluación en un caso de uso de *sobres digitales con capacidades de búsqueda (SDB)*,

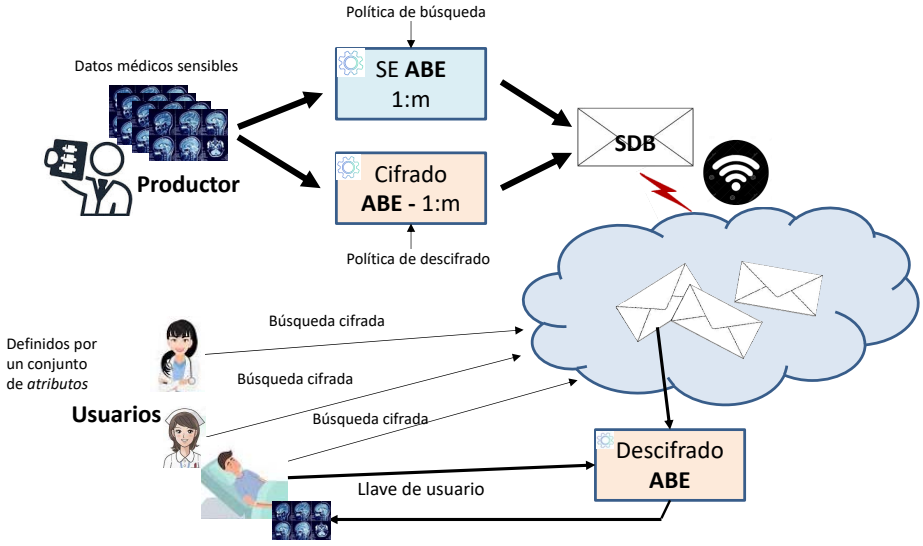


Figura 1: Vista general del concepto de sobres digitales con capacidades de búsqueda y su aplicación en el dominio de *big data* en salud.

los cuales son abstracciones fundamentadas en el cifrado de datos no convencional, llamado cifrado basado en atributos (ABE) [17]. Mediante ABE y bajo el concepto de SDB abordamos los tres problemas previamente descritos. En las siguientes secciones se darán detalles tanto de ABE como de su implementación con capacidades de búsqueda (SE-ABE). Por ahora, el enfoque de solución y concepto de SDB se muestra en la Figura 1. Con base en ello, el proceso de aseguramiento de datos sensibles en salud mediante SDB se lleva a cabo de la siguiente forma:

1. El productor de datos, que generalmente es un médico, especialista en salud, o incluso algún dispositivo médico, genera datos sensibles, como radiografías;
2. Estos datos en algún momento serán requeridos por otro médico o profesional de la salud, incluso por el mismo paciente; al ser las radiografías en nuestro ejemplo datos sensibles, éstas no pueden almacenarse en un disco o memoria, o enviarse mediante correo o algún servicio de mensajería a través de un teléfono inteligente; los datos deben asegurarse y almacenarse en un medio desde el cual después dichos datos puedan recuperarse; para ello, el productor de los datos debe ejecutar dos macro-procesos: *i*) el algoritmo de cifrado ABE, el cual requiere de una política de control de acceso que, en esencia, determina qué usuarios podrán descifrar esos datos más adelante; *ii*) el algoritmo de búsquedas cifradas SE-ABE, el cual también requiere de una política de control de acceso, pero que determina qué usuarios podrán consultar los datos cifrados mediante *tokens* igualmente cifrados;

3. Tanto el cifrado ABE como parte de la búsqueda cifrada SE-ABE se ejecutan del lado del productor de datos; el resultado de cada proceso es, por un lado, los datos cifrados y, por otro, un índice seguro de búsqueda sobre los datos cifrados; todo esto es lo que conforma el SDB;
4. Una vez creado, el SDB es enviado a un repositorio (la nube), de donde más adelante los usuarios autorizados podrán hacer consultas (si sus atributos satisfacen la política usada en SE-ABE), recuperar la información de interés (cifrada) y acceder a ella descifrándola usando su llave de usuario estrechamente relacionada con los atributos que le describen (médico, enfermera, paciente y demás datos relacionados con dicho rol);
5. Un usuario de los datos está descrito por los atributos que lo caracterizan; por ejemplo, un profesional de la salud puede tener atributos como su especialidad, nivel jerárquico en la organización donde labora, datos del lugar donde labora que describen su pertenencia a dicha organización, datos personales y cualquier otro que sea relevante para el propósito; con base en estos atributos, cada usuario cuenta con una llave, la cual es intransferible y necesaria para crear los *tokens* de búsqueda y para descifrar los datos recuperados desde el repositorio de datos cifrados.

Nuestra construcción de SDB se realiza sobre emparejamientos bilineales asimétricos, los cuales son estructuras matemáticas en el dominio de la teoría de grupos y campos finitos. Con ello, ABE no solo es lo suficientemente seguro para preservar la confidencialidad de los datos, sino también para garantizar el control de acceso a éstos únicamente a entidades específicas mediante el cifrado de uno a muchos. Asimismo, se habilitan las búsquedas cifradas también bajo el concepto de cifrado basado en atributos, reutilizando las estructuras algebraicas de esta técnica criptográfica. El cifrado de datos en SDB consta de dos capas de cifrado a partir de las cuales se toma ventaja a la par de la eficacia y del elevado nivel de seguridad provisto por ABE y de la eficiencia de los cifradores simétricos [18]. Los datos sensibles son cifrados mediante llaves de sesión de un cifrador simétrico, mientras que las llaves de sesión se cifran a partir de ABE. De esta forma, es posible preservar la confidencialidad de los datos y decidir selectivamente los usuarios autorizados para acceder y consumir determinados conjuntos de datos.

Este capítulo está organizado de la siguiente manera: en la Sección 2 presentamos los conceptos más relevantes para la definición de los SDB; en la Sección 3 se presentan los detalles de diseño de sobres digitales, mientras que en la Sección 4 se presenta el diseño de los SDB; en la Sección 5 se describen las estrategias de eficiencia y paralelismo de los SDB, mientras que en la Sección 6 se detalla la validación de la construcción de SDB en el dominio de la salud, siendo un componente principal en el despliegue de un servicio que permite el aseguramiento de datos médicos desde su producción hasta su consumo [4]; en la Sección 7 se discuten, desde la perspectiva de este trabajo, los retos para proteger los datos en el sector salud; finalmente, en la Sección 8 se presentan las conclusiones.

2. Antecedentes

Día con día, usuarios e instituciones generan y recolectan grandes cantidades de datos derivados de actividades cotidianas, tales como compras en línea, realización de trámites, transacciones bancarias, entre muchas otras. Gran parte de estos datos se encuentra disponible de forma pública en sitios web o redes sociales, por citar algunos ejemplos. Sin embargo, existen datos que son sensibles y requieren protección ante el acceso no autorizado por parte de terceros [6]. Ejemplos de datos sensibles son los registros médicos o financieros, números de cuenta bancarios, datos de identificación personal o número de seguridad social; planes de adquisición, información personal de clientes, datos financieros o derechos de propiedad intelectual. De esta forma, a medida que se generan y distribuyen grandes cantidades de datos, la protección de éstos se vuelve indispensable para sus propietarios, ya sean individuales o grandes empresas e instituciones.

La importancia de la seguridad de la información radica en la protección de los datos y sistemas que los producen o utilizan del daño, uso, divulgación o destrucción no autorizados [11]. Con el volumen y variedad de datos, así como la velocidad con la que éstos son generados por los usuarios y las operaciones diarias del negocio, la confidencialidad, integridad y disponibilidad de los datos son esenciales para la seguridad de la información [6]. Uno de sus principios fundamentales es la confidencialidad, la cual garantiza la privacidad de los datos al restringir el acceso a éstos través del cifrado de su contenido. Para ello, se apoya, además, en mecanismos de autenticación o concesión de niveles de privilegios, permitiendo que solamente personas autorizadas puedan ver o manipular datos. De esta manera se evita que entidades no autorizadas puedan derivar u obtener información a partir de dichos datos [6], [7].

2.1. Criptografía

Uno de los métodos más utilizados para asegurar la confidencialidad de los datos es el cifrado. La criptografía se encarga de implementar el cifrado realizando transformaciones a los datos de manera que, al almacenarlos y transmitirlos, solo los destinatarios autorizados puedan accederlos y procesarlos [10]. Dichas transformaciones se llevan a cabo del lado de los propietarios de datos, convirtiendo información legible en texto incomprensible, y de los destinatarios, aplicando el proceso inverso de descifrado para obtener los datos originales [12], [10]. Los métodos criptográficos modernos emplean algoritmos seguros desde el punto de vista computacional con la finalidad de que la información protegida no pueda ser comprometida fácilmente. Dicho objetivo se logra mediante mecanismos como el cifrado de los datos o de los canales de comunicación durante su transmisión, así como la generación de códigos de autenticación de mensajes y firmas digitales [12], [7].

Los procesos de cifrado y descifrado requieren dos componentes elementales: un algoritmo criptográfico, o cifrador, y una llave. El algortimo consiste en la aplicación de funciones matemáticas que, en conjunto con una llave que se utiliza

como parámetro de entrada durante el procedimiento, realizan las transformaciones pertinentes a los datos [12]. Al llevar a cabo el cifrado de los datos, éstos se protegen al convertir su forma legible en texto cifrado que resulta incomprensible. En contraste, el descifrado constituye el proceso opuesto, a partir del cual se obtienen los datos sensibles al remover la protección suministrada a través del cifrado [7]. Por otra parte, la llave o clave criptográfica es un componente indispensable de cualquier algoritmo de cifrado. Usualmente, estas claves se generan de manera aleatoria previa al cifrado de los datos, aunque éstas también pueden ser especificadas por el usuario [12], [11].

Cabe destacar que la ventaja del uso de un algoritmo sobre otro radica en la eficacia de la generación y administración de las llaves utilizadas para los procesos de cifrado y descifrado. Cuanta mayor dificultad asociada a la clave criptográfica, mayor seguridad será capaz de brindar el algoritmo; no obstante, su ejecución se vuelve más compleja. En consecuencia, la seguridad del cifrado reside sustancialmente en el secreto de las claves, no en el algoritmo. En la actualidad, existen diversos algoritmos de cifrado y, debido a que éstos son de acceso público, las claves criptográficas que dichos algoritmos utilizan son las que garantizan la discreción de los datos [7]. Son diversos los sistemas criptográficos que han presentado fallas debido a errores en sus procedimientos de administración de llaves. En la práctica, la mayoría de los ataques implican vulnerar el sistema de gestión de claves, en lugar del algoritmo criptográfico en sí. Es por ello que la gestión de claves constituye la parte más difícil de abordar al momento de diseñar un sistema criptográfico.

Cada método de cifrado utiliza un algoritmo específico (P, como ha sido descrito en la Ecuación 1) que se compone de una serie de pasos bien definidos, generalmente estandarizados, utilizados para cifrar y descifrar los datos. Existen diversos métodos para crear texto cifrado, siendo los más antiguos la transposición o sustitución de caracteres y, el más reciente, la combinación de datos con claves secretas. Si bien los algoritmos de cifrado contemporáneos emplean técnicas más robustas basadas en problemas matemáticos, no hay un algoritmo que resulte idóneo para cualquier caso de aplicación. La adopción de éste dependerá del nivel de sensibilidad y cantidad de información que se requiera proteger, así como de los inconvenientes que se pretendan mitigar. Asimismo, la forma de almacenamiento o de transmisión de los datos y los recursos computacionales con los que se cuente determinarán, en gran medida, la opción que habrá de elegirse para llevar a cabo esta tarea [22], [7]. Según la manera en que se gestionen las llaves, los algoritmos criptográficos se pueden dividir en simétricos y asimétricos.

2.2. Criptografía de Clave Privada: Cifradores Simétricos

Los algoritmos simétricos se caracterizan por hacer uso de una cantidad menor de recursos computacionales que su contraparte asimétrica. Esto se debe a que en el cifrado simétrico se utiliza una misma llave para realizar el cifrado y descifrado de los datos, tal como se observa en la Figura 2. A esta llave se le denomina clave secreta o previamente compartida, ya que el emisor y recep-

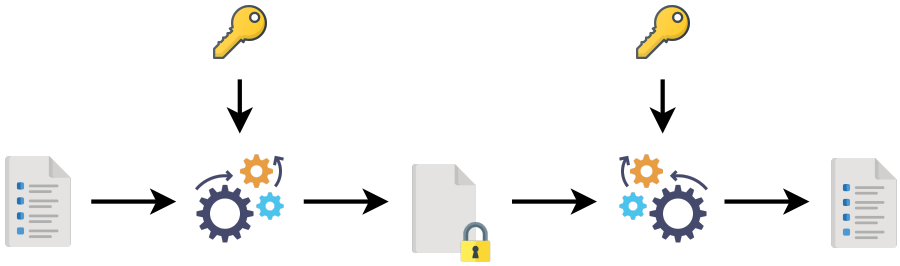


Figura 2: Flujo de operaciones en el cifrado simétrico.

tor deben conocerla antes de que inicie el proceso de cifrado [7], [12]. Dado un mensaje, la clave secreta sirve como parámetro de entrada para el algoritmo de cifrado que aplica las transformaciones necesarias para producir como salida un texto cifrado. Cabe destacar que dicho mensaje se procesa a nivel de arreglos de bytes, por lo que éste puede representar desde una cadena de caracteres hasta un archivo de cualquier extensión. Por el contrario, el algoritmo de descifrado recibe como entrada el texto cifrado, así como la misma clave previamente compartida, y produce como resultado el texto plano del mensaje original.

Los algoritmos simétricos pueden dividirse en cifradores por bloque o por flujo en función de la cantidad de datos de entrada que manejan. Es decir, la diferencia entre ambos recae en la forma de realizar el agrupamiento de bits para los procesos de cifrado y descifrado. Los algoritmos de cifrado por bloque dividen los datos de entrada en bloques de tamaño fijo, usualmente de 64 ó 128 bits, y posteriormente realizan el procesamiento de dichos bloques. En cambio, los cifradores por flujo procesan los datos de entrada conforme éstos se van recibiendo, esto es, un byte o un bit a la vez [7]. De acuerdo con Barker [2], el estándar de cifrado avanzado (AES) es el algoritmo recomendado en la actualidad por el Instituto Nacional de Estándares y Tecnología (NIST) para el cifrado-descifrado de datos. AES es un cifrador por bloque desarrollado para reemplazar al ya obsoleto estándar de cifrado de datos (DES), por lo que constituye el algoritmo más utilizado en la actualidad. Este algoritmo procesa datos en bloques de 128 bits utilizando claves de 128, 192 ó 256 bits, con lo cual se considera que es capaz de proveer niveles de seguridad válidos más allá del año 2030.

2.3. Criptografía de Clave Pública: Cifradores Asimétricos

Para el cifrado y descifrado, los algoritmos asimétricos utilizan un par de claves –una pública y otra privada– matemáticamente relacionadas entre sí, lo cual implica que estos algoritmos sean más complejos que su contraparte simétrica. Si bien existe una relación matemática entre las claves, no es posible obtener la llave privada a partir de la llave pública debido a la complejidad de las operaciones involucradas en la generación del par de claves. Es por ello que estos algoritmos requieren una mayor cantidad de recursos computacionales y su pro-

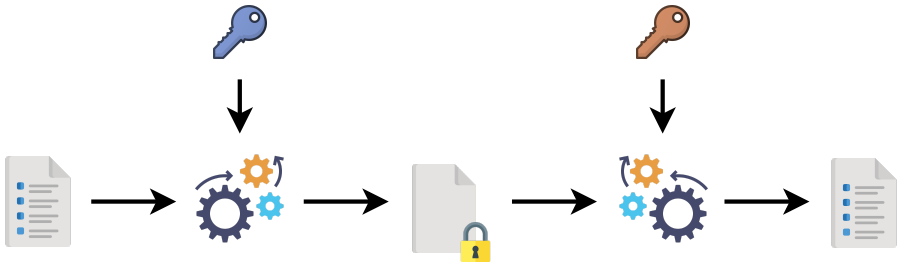


Figura 3: Flujo de operaciones en el cifrado asimétrico.

ceso de ejecución es más lento. En el ejemplo expuesto en la Figura 3, en un sistema de cifrado de clave pública el emisor cifra los mensajes utilizando la clave pública de un receptor en particular [12], [11]. De esta forma, únicamente la clave privada asociada a la clave pública usada para el cifrado puede descifrar los datos. Debido a ésto, no es necesario que ambas partes hagan uso de una llave previamente compartida para intercambiar mensajes de manera segura.

Al igual que en la criptografía de clave privada, los algoritmos asimétricos son de acceso público, por lo que es posible conocer cómo trabajan de manera general. Sin embargo, sus operaciones se basan en problemas matemáticos complejos pero bien conocidos, como la factorización de enteros y los logaritmos discretos. Dado lo anterior, la seguridad de este tipo de algoritmos recae en el par de claves de las cuales se hace uso en los procesos de cifrado y descifrado [7]. Al utilizar funciones del álgebra abstracta en lugar de números reales se vuelve poco viable la búsqueda de la clave privada asociada a una clave pública, incluso si se conoce con qué algoritmo se crearon ambas, dado el tiempo computacional requerido y el elevado costo asociado a dicho procedimiento [12]. Rivest-Shamir-Adleman (RSA), Diffie-Hellman (DH), ElGamal y la criptografía de curva elíptica (ECC) son ejemplos de algoritmos asimétricos utilizados hoy en día.

Como se puede observar, existen evidentes diferencias entre los algoritmos simétricos y asimétricos.

Los primeros son más eficientes en cuanto al tiempo de procesamiento requerido y la cantidad de datos que pueden manejar, pero la gestión de claves es mucho más difícil por el uso de claves compartidas. Otra de las grandes diferencias entre ambos es el tamaño de las llaves, lo cual determina su susceptibilidad a ataques por fuerza bruta. Al ser las claves asimétricas de mayor tamaño en comparación con las claves simétricas –1024 bits contra 128 bits, respectivamente–, el rango de valores posibles es también mucho mayor, lo cual hace inviables este tipo de ataques [11]. En la práctica, ambos esquemas se utilizan en conjunto: el cifrado simétrico se utiliza para garantizar la confidencialidad de grandes volúmenes de datos y la criptografía de clave pública para el intercambio de las claves secretas.

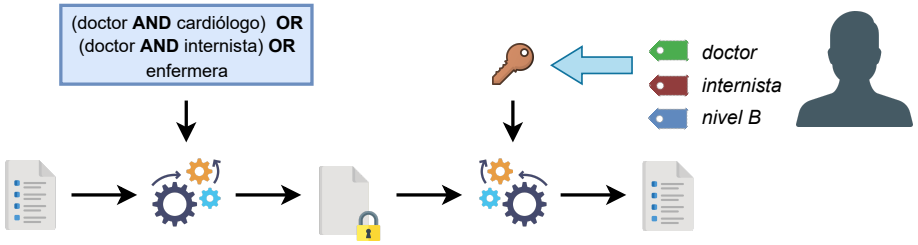


Figura 4: Flujo de operaciones en el cifrado basado en atributos.

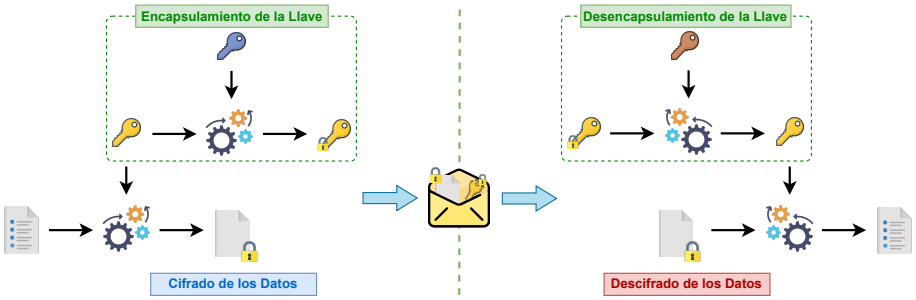
2.4. Cifrado Basado en Atributos

El cifrado basado en atributos (ABE) es una técnica de la criptografía de clave pública que tiene su fundamento en un mecanismo de control de acceso en el que una entidad se identifica a partir de un conjunto de atributos descriptivos. Éste permite compartir datos de forma segura con múltiples usuarios, a la vez que ofrece una gran flexibilidad de gestión de acceso a los datos [17], [1]. En lugar de utilizar las tradicionales claves públicas o privadas, en ABE los datos se cifran mediante la especificación de los atributos que un potencial usuario debe poseer para poder descifrar un mensaje utilizando su clave secreta, tal como se muestra en la Figura 4. Dichos atributos se especifican en estructuras denominadas políticas de control de acceso, las cuales establecen las reglas de acceso a los datos de los propietarios mediante compuertas lógicas (*AND*, *OR*) o de tipo umbral (*k-of-n*) [14].

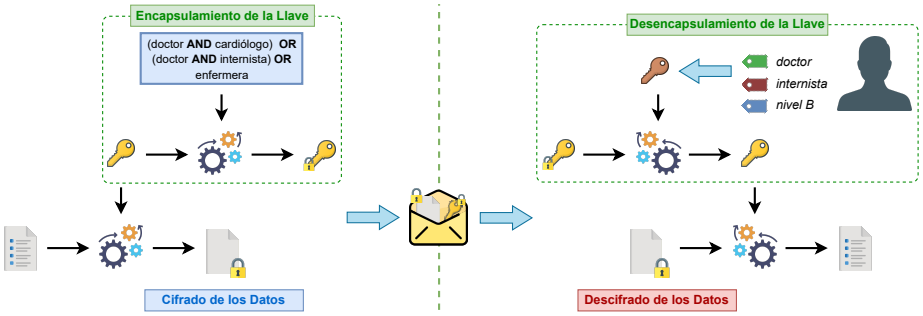
Cabe mencionar que, a diferencia de otros algoritmos de clave pública, ABE es un esquema de cifrado de muchos a muchos, por lo que los propietarios de datos no tiene que conocer de antemano a todos los posibles usuarios. Dado que los algoritmos asimétricos tradicionales utilizan un par de claves relacionadas matemáticamente, un mismo mensaje se tiene que cifrar tantas veces como destinatarios existan para éste. Por el contrario, puesto que en ABE un atributo puede ser común a múltiples usuarios, el cifrado mediante políticas de acceso permite abarcar un mayor número de destinatarios. Esta característica representa una de las principales ventajas de este esquema, ya que permite un control de acceso de grano fino, sin incurrir en sobrecargas de almacenamiento y comunicación asociadas a algoritmos como RSA [19]. De este modo, ABE resulta más adecuado para escenarios de almacenamiento y compartición de datos en la nube, ya que los datos de los propietarios permanecen confidenciales incluso en entornos poco confiables.

3. Sobres Digitales

Como se mencionó anteriormente, la confidencialidad de datos sensibles se logra a partir del cifrado, y éste puede realizarse mediante algoritmos simétricos o asimétricos. Ambos realizan transformaciones a los datos de modo que solamente aquellos destinatarios que posean la correspondiente llave de descifrado



(a) Criptografía de clave pública tradicional.



(b) Cifrado basado en atributos.

Figura 5: Flujo de operaciones en la creación y uso de sobres digitales.

puedan acceder a dichos datos. Sin embargo, estos dos tipos de cifradores conllevan desventajas que dificultan su uso de forma aislada; por ello, éstos suelen emplearse de forma conjunta en aplicaciones prácticas. La ventaja de los algoritmos simétricos sobre los asimétricos es la capacidad de cifrar una gran cantidad de datos eficientemente en términos de tiempos de respuesta. No obstante, éstos implican un problema de distribución y gestión de llaves debido a que la llave usada para cifrar es la misma requerida por el proceso de descifrado. Así, es necesario que los propietarios de datos compartan las claves de descifrado con los destinatarios de dichos datos a través de canales de comunicación seguros, algo que no es posible garantizar en todos los casos.

En este sentido, una forma de sortear el problema de compartición de claves es mediante el uso de una técnica criptográfica denominada *sobre digital*. Un sobre digital se define como un objeto criptográfico que consta de dos capas de cifrado a partir de las cuales se transporta y distribuye una llave de sesión de forma segura. Mediante éstos, es posible tomar ventaja simultáneamente tanto de la criptografía de clave pública, como de la criptografía de clave privada. Como se puede observar en la Figura 5a, los datos sensibles son cifrados mediante llaves de sesión de cifradores simétricos, mientras que dichas llaves de sesión se cifran o encapsulan a partir de criptografía de clave pública. Es decir, los datos se cifran y descifran con una misma llave simétrica, que a su vez se cifra utilizando la clave

pública del destinatario de los datos y éste los descifra usando su clave privada. De esta forma es posible preservar la confidencialidad de una gran cantidad de datos en un tiempo razonable, mientras que es posible compartir las llaves de descifrado con destinatarios específicos, aun utilizando canales de comunicación inseguros.

Además de proporcionar mayor robustez contra ataques, los algoritmos asimétricos eluden el problema de compartición de llaves al utilizar un par de claves relacionadas matemáticamente. No obstante, dada dicha relación entre llaves, es necesario conocer a priori a los potenciales usuarios de un mismo conjunto de datos. Por ello, una forma de abordar esta problemática en el contexto de los sobres digitales implica la remoción del algoritmo asimétrico empleado y la incorporación en su lugar del cifrado basado en atributos, tal como se muestra en la Figura 5b. De esta manera, ABE permite compartir datos de forma segura con múltiples usuarios, incluidos aquellos no definidos a priori. Solamente aquellos usuarios que posean un conjunto de atributos que satisfaga de forma criptográfica la política de control de acceso definida previo al cifrado podrán acceder a los datos en texto plano. Es decir, únicamente quienes cumplan con los criterios establecidos en la política de acceso podrán acceder a la llave de sesión y, con ella, a los datos sensibles.

De esta manera, al emplear sobres digitales en conjunto con ABE desaparece la necesidad de implementar mecanismos adicionales de gestión de llaves. Lo anterior, considerando que los atributos permiten describir las características de los usuarios, así como sus inherentes derechos de acceso. De esta forma, se evitan sobrecargas de cómputo, resulta poco significativo si la transmisión de datos se realiza mediante canales de comunicación seguros o no, y se impone un control de acceso de grano fino. Incluso si un sobre es filtrado pero su portador no cuenta con los atributos que satisfacen la política utilizada en la creación de dicho sobre, éste no será capaz de acceder al contenido legible del sobre digital. Cabe destacar que, a partir de los atributos que posean los usuarios, una autoridad de confianza (TA) se encarga de generarle a cada usuario su correspondiente llave secreta, la cual permite corroborar si éste satisface la política utilizada en el cifrado. Además, la TA tiene la facultad de implementar mecanismos de revocación de acceso para el caso de aquellos usuarios que dejen de pertenecer a la organización donde se gestionan los datos sensibles o aquellos que hagan uso indebido de los mismos.

4. Sobres Digitales con Capacidades de Búsqueda

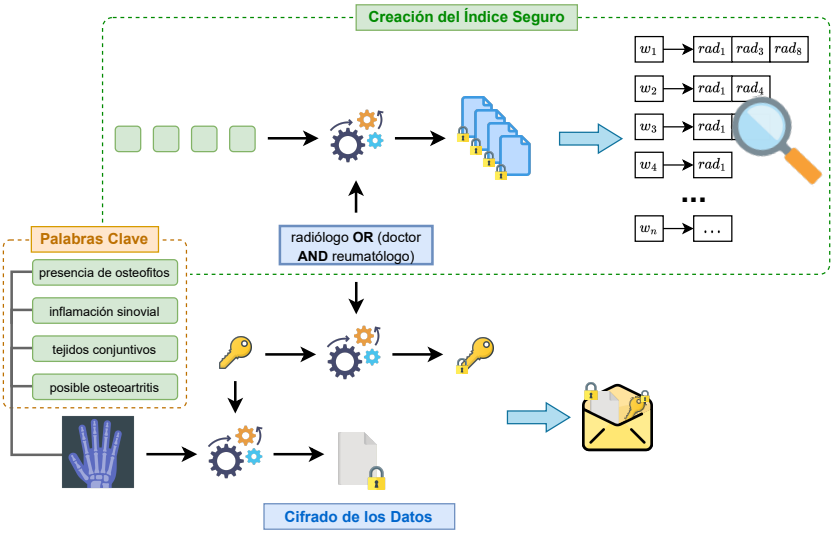
El cifrado de datos permite garantizar la confidencialidad de datos sensibles en entornos poco confiables, como en el caso de escenarios de almacenamiento en la nube. Asimismo, el almacenamiento en la nube facilita el acceso conveniente a los datos y su compartición con múltiples usuarios, proveyendo, además, capacidades de búsqueda y recuperación de información. Sin embargo, el hecho de que los propietarios cifren sus datos previo a externalizarlos a la nube introduce dos grandes problemas. En primer lugar, la compartición se vuelve una tarea

compleja que implica que los propietarios gestionen los mecanismos de control de acceso hacia sus datos. Dicha problemática se puede abordar a través del uso de ABE, el cual ofrece una gestión flexible mediante controles de acceso de grano fino a la vez que garantiza la confidencialidad de los datos. En segundo lugar, las capacidades de búsqueda del proveedor del servicio de almacenamiento no se pueden aprovechar debido a que los datos se encuentran en formato ininteligible.

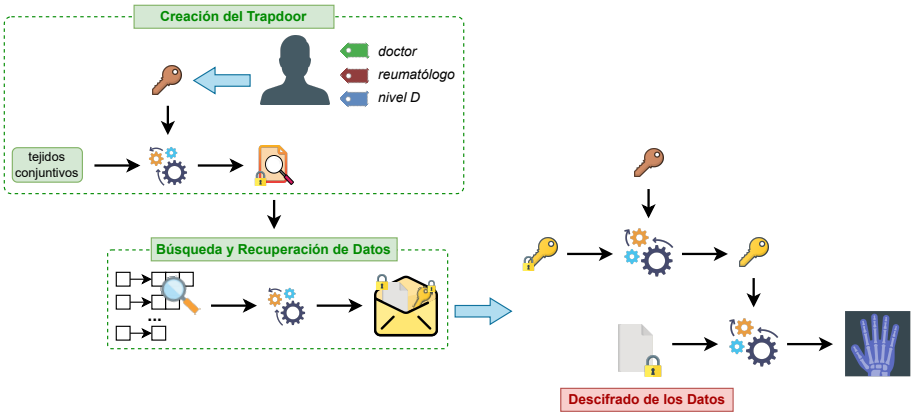
Si bien los usuarios podrían descargar todos los datos (cifrados), descifrarlos y aplicar localmente algoritmos de búsqueda y recuperación tradicionales, este enfoque es totalmente inviable en la práctica por varias razones. Por ejemplo, se introducen sobrecargas de comunicación innecesarias al descargar todo un conjunto de datos que, en el peor de los casos, pudiera no contener información relevante. Además, se generan sobrecargas de procesamiento donde, al poseer recursos heterogéneos, no todos los dispositivos pueden ejecutar procedimientos exhaustivos de búsqueda. En este contexto, surge *Searchable Encryption* (SE), una técnica criptográfica que permite realizar búsquedas sobre datos cifrados. Su objetivo es mantener la confidencialidad de los datos mientras el proveedor del servicio de almacenamiento es capaz de preservar sus capacidades de búsqueda [5]. SE ha sido implementado mediante tres enfoques principales, siendo el cifrado basado en atributos con capacidades de búsqueda (ABSE) el más adecuado para escenarios de almacenamiento y compartición de datos cifrados.

ABSE se apoya en la creación de un índice seguro que contiene palabras clave representativas del contenido o características de los datos sensibles y, a partir del cual, posteriormente se realizan las búsquedas [1]. Al ser un enfoque basado en atributos, ABSE opera de forma muy similar a ABE: se emplean políticas para establecer reglas de acceso y atributos para describir a los usuarios y, por ende, sus restricciones de acceso. Tras identificar las palabras clave que describen el contenido de los datos, éstas se cifran una sola vez mediante una política de acceso, definida sobre un conjunto de atributos y, a partir de ellas, se construye el índice seguro. Una vez creado éste, tanto los datos sensibles como su correspondiente llave de sesión son cifrados, produciendo el sobre digital que habrá de enviarse en conjunto con el índice seguro para su almacenamiento en la nube [9]. Cabe resaltar que la política de acceso utilizada para cifrar las palabras clave puede ser la misma o una política diferente a la usada para cifrar la llave de sesión, dependiendo de las necesidades de acceso que caractericen a los datos. Este proceso se ilustra en la Figura 6a.

Para realizar las búsquedas, el índice seguro es consultado por el proveedor del servicio de almacenamiento dado un *token* cifrado, denominado trampilla de búsqueda o *trapdoor*, creado por el usuario que solicita una búsqueda. Al igual que en ABE, cada usuario posee una llave secreta que se genera con base en el conjunto de atributos que lo caracterizan. De este modo, solo los usuarios que poseen el conjunto de atributos adecuados (dada una política) pueden buscar y recuperar datos de interés [1], [13]. Derivado de una necesidad de información, a partir de la clave secreta de usuario se genera una representación cifrada de la consulta del usuario, la cual permite realizar la búsqueda en el índice seguro [20]. Por ello, es importante señalar que el proveedor de servicio no es capaz de



(a) Creación del índice seguro y cifrado.



(b) Creación del *trapdoor*, búsqueda y descifrado.

Figura 6: Flujo de operaciones en la creación y uso de sobres digitales con capacidades de búsqueda.

derivar información, saber qué está buscando o el contenido de los resultados que encuentra dado un trapdoor en particular.

Si los atributos del potencial usuario de los datos satisfacen la política de cifrado y si existen resultados para su consulta, se retornan los sobres digitales correspondientes. Finalmente, el usuario podrá descifrar la llave de sesión mediante su clave secreta y los datos por medio de dicha llave de sesión, tal como se muestra en la Figura 6b [9], [20]. De esta forma se garantiza confidencialidad y control de acceso, así como la capacidad de compartir datos de forma segura con múltiples usuarios al incorporar ABE en el contexto de los sobres digitales.

Además, se preservan las capacidades de búsqueda y recuperación de información del proveedor del servicio de almacenamiento al hacer uso de esta técnica en conjunto con el cifrado con capacidades de búsqueda.

5. Eficiencia y Seguridad de Sobres Digitales

Crear y abrir SDBs requiere suficiente poder de cómputo. La complejidad en tiempo de ejecución y demanda de recursos de cómputo está asociada a los algoritmos criptográficos para cifrar los datos con el cifrador simétrico y para proteger la llave de sesión de dicho cifrador simétrico mediante el cifrado basado en atributos. Esta complejidad queda determinada por:

- El tamaño de los datos a cifrar, que impacta directamente en la complejidad en tiempo para el cifrador simétrico;
- El nivel de seguridad, que impacta directamente en el número de operaciones y la longitud de los operandos del cifrado basado en atributos.

El problema de eficiencia en SDBs se aborda a través del uso de patrones de paralelismo. El problema de la seguridad en SDBs se aborda mediante construcciones basadas en emparejamientos asimétricos. Se consideran dos esquemas de paralelismo: *i*) *pipeline*; y *ii*) *overlapped*. En el esquema llamado *pipeline* se despliegan dos patrones de paralelismo diferentes: el patrón *pipe & filters* en combinación con el patrón conocido como *manejador/trabajador*. De esta manera, el patrón *pipe & filters* se encarga de organizar el sistema en tuberías y el patrón manejador/trabajador se encarga de desplegar dichas tuberías como trabajadores ejecutados en paralelo. Este esquema se encarga de cifrar y descifrar

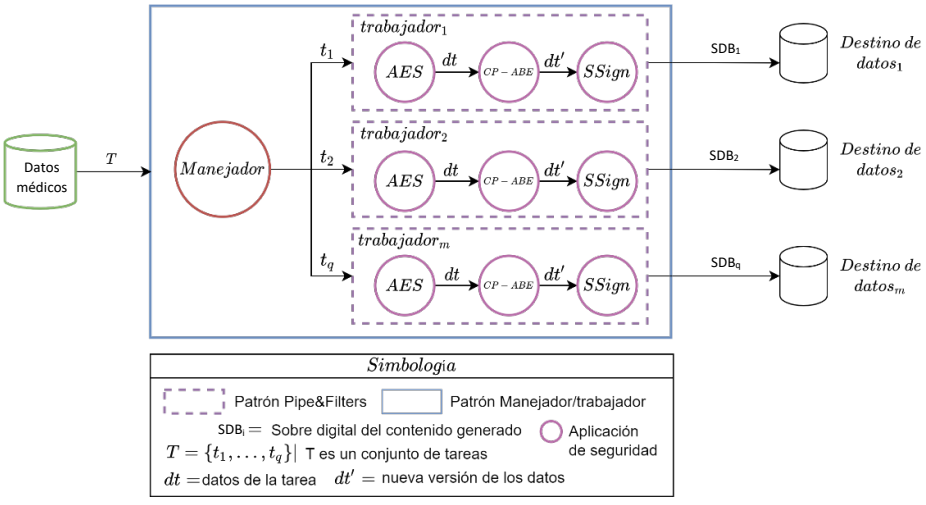


Figura 7: Representación conceptual de un esquema *pipeline*.

conjuntos pequeños de datos dividiendo las tareas entre el número de trabajadores disponibles desplegados en el sistema, lo cual permite reducir el tiempo de ejecución de los procesos de cifrado, los cuales dependen del tamaño de los datos, más que del nivel de seguridad.

La Figura 7 muestra la representación conceptual del ejemplo de un esquema *pipeline*, el cual cuenta con un patrón *pipe & filters* que incluye las aplicaciones de AES (como cifrador simétrico), CP-ABE (como cifrador basado en atributos) y SSign (para firma digital, basado en identidad). En este esquema, cada tubería es clonada en q trabajadores para eficientizar el procesamiento de los datos mediante la distribución de tareas a través de un trabajador dedicado. Una vez que los datos, expresados como archivos t_i en un repositorio de entrada, son procesado a través de toda la tubería por un trabajador dedicado, éstos se encapsulan para dar origen al SDB_i correspondiente, que puede ser enviado a su destino.

Por otro lado, el esquema *overlapped* permite el acoplamiento de sistemas independientes para que se ejecuten de forma suprapuesta (mediante el patrón *fork/join*), y el acoplamiento en forma de tubería para aquellos sistemas que cuenten con algún tipo de dependencia. Este esquema permite que los procesos asociados con la creación de SDBs se ejecuten en forma de una tubería y se gestionen como si fueran trabajadores en un patrón manejador/trabajador, permitiendo que la ejecución de tareas se realice de forma paralela. El esquema *overlapped* fue diseñado para cifrar y descifrar grandes conjuntos de datos de forma paralela.

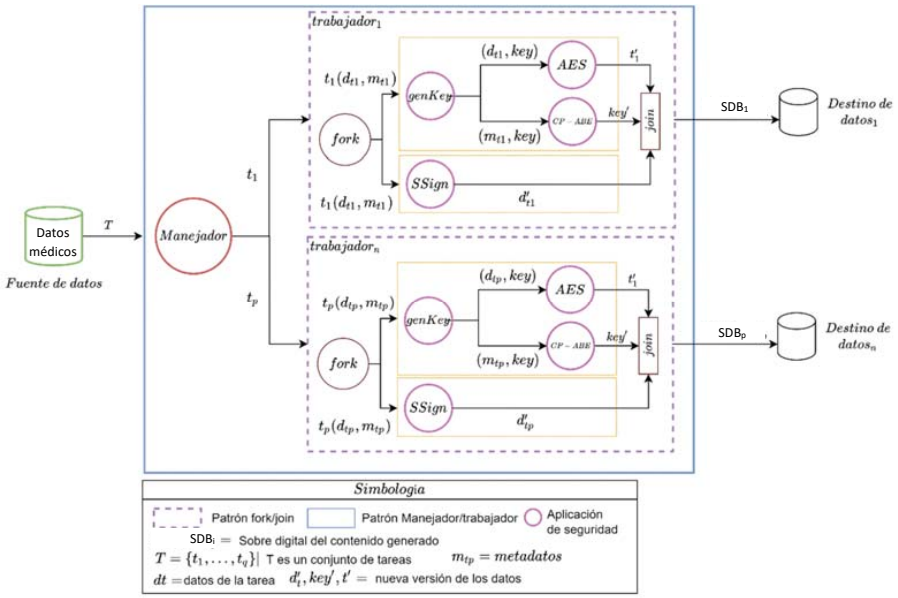


Figura 8: Representación conceptual de un esquema *overlapped*.

La Figura 8 muestra la representación conceptual de un ejemplo de un esquema *overlapped*, en donde un conjunto de tareas es extraído desde una fuente de datos y, posteriormente, éstas son distribuidas a los trabajadores a través de un manejador. Cada trabajador contiene un patrón *fork/join*, el cual permite la ejecución de tareas de forma suprapuesta (en este caso, la ejecución de la generación de llaves al mismo tiempo que la ejecución de la firma digital de los contenidos). Esto permite ejecutar dos tuberías al mismo tiempo y, una vez que la ejecución de ambas tuberías ha terminado, el contenido procesado es integrado en un sobre digital y, posteriormente, enviado a un destino (ya sea para su consumo, compartición o procesamiento en un entorno distinto).

6. Sobres Digitales para Ciencia de Datos

Hasta nuestro conocimiento, los SDBs han sido explorados y propuestos como tal, por primera vez, en el grupo de investigación del Cinvestav Unidad Tamaulipas [9]. De igual forma, y en el marco del programa de apoyo a la investigación en salud PRONACES - Salud, PRONAII Ciencia de Datos en Salud, en el proyecto *"Plataforma tecnológica para la gestión, aseguramiento, intercambio y preservación de grandes volúmenes de datos en salud y construcción de un repositorio nacional de servicios de análisis de datos de salud"* [15], se han implementado los sobres digitales para garantizar la seguridad, de extremo a extremo, de los datos médicos durante su ciclo de vida, que incluye:

1. **Creación**: Los datos se originan en un dispositivo médico (como radiografías, por ejemplo) o sistema de información (expediente clínico electrónico).
2. **Almacenamiento**: Una vez creados, los datos se encapsulan en SDBs y se almacenan en repositorios, locales o externos. En el caso de usar medios de almacenamiento externo, la comunicación desde el origen al destino se realiza comúnmente por un medio público o inseguro. Las propiedades de seguridad inherentes de los SDBs permiten emplear, incluso, canales de comunicación inseguros.
3. **Uso**: Los datos que ya se encuentran en un repositorio pueden ser consumidos o consultados por usuarios autorizados mediante técnicas de búsqueda, recuperación y acceso a SDB. Las características inherentes a los SDBs permiten realizar estas operaciones que habilitan a los usuarios autorizados (médicos, especialistas, profesionales de la salud, u otros dispositivos o sistemas) acceder a los datos de manera segura.

La Figura 9 describe de manera gráfica el ciclo de vida descrito previamente. Los datos se crean en el ámbito de una organización A (hospital, unidad médica familiar, consultorio, laboratorio) y es ahí donde se aseguran mediante la creación del SDB. Más adelante, desde el repositorio donde se encuentre dicho SDB, los usuarios (especialistas de la salud, sistemas, dispositivos) pueden acceder a él y únicamente aquellos con los atributos necesarios podrán abrir el SDB y acceder a los datos en claro. Durante la creación y apertura de SDB,

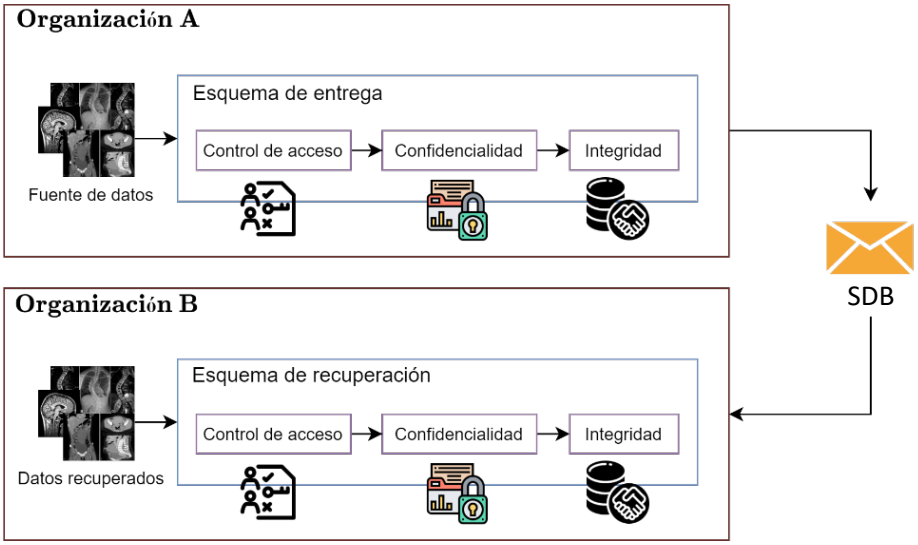


Figura 9: Creación y acceso (apertura) de sobres digitales buscables (SDB) en salud.

los esquemas de paralelismo permiten la viabilidad de implantar este concepto, puesto que los datos médicos, como las tomografías, generalmente ocupan una cantidad considerable de almacenamiento y, bajo un escenario de *big data*, la complejidad en tiempo incrementa considerablemente. El transporte seguro de los datos sensibles médicos, posible mediante los SDB, permite la distribución de datos sensibles de forma segura, lo cual es un requerimiento en las normas oficiales para tratamiento de datos médicos. Los usuarios finales de los datos, que los recuperan mediante operaciones de descifrado, podrán usarlos en los procesos correspondientes para su análisis y obtención de conocimiento útil mediante técnicas de ciencia de datos.

7. Retos y Perspectivas para Proteger Datos en el Sector Salud

Como es sabido, los datos sensibles generalmente demandan servicios de seguridad y de privacidad. Éste es un requerimiento impuesto, incluso, por regulaciones y leyes que varían en cada país. En México, la norma oficial NOM-024-SSA3-2012 establece los objetivos funcionales y funcionalidades que deberían observar los productos de Sistemas de Expediente Clínico Electrónico para garantizar la interoperabilidad, procesamiento, interpretación, confidencialidad, seguridad, así como uso de estándares y catálogos de la información de los registros electrónicos en salud. En este sentido, los SDBs son mecanismos que permiten lograr objetivos en materia de seguridad de datos y que permiten dar cumplimiento a los requerimientos de seguridad de éstos. Específicamente, la confidencialidad y

control de acceso son esenciales en el manejo, acceso e intercambio de datos en salud, que se consiguen con el uso de SDB mediante las técnicas criptográficas que incorporan.

El cifrado usado en los SDBs es de dos tipos: simétrico y basado en atributos. Ambos enfoques actualmente son suficientemente seguros. El cifrado simétrico sustenta su seguridad en que el trabajo de un atacante para vulnerarlo es exponencial respecto al tamaño de las llaves usadas. En la actualidad, se considera que una llave de 128 bits es imposible de atacar (en la práctica), dado que el costo computacional para lograrlo es de 2^{128} operaciones. Por lo anterior, a una computadora convencional le tomaría centenas de años realizar este trabajo. En el caso del cifrado basado en atributos, las llaves utilizadas son de un tamaño de al menos de 3×256 bits. En este caso, el costo para un atacante depende de la dificultad para resolver un problema matemático con un costo asociado de 2^{256} operaciones en una computadora convencional, pero la llave involucra al menos tres componentes de 256 bits.

Sin embargo, desde hace algunos años se vienen teniendo avances significativos en cómputo cuántico. Incluso, algunas empresas como Google han declarado haber alcanzado ya la supremacía cuántica, esto es, haber logrado diseñar una computadora que resuelve problemas que una computadora convencional no habría podido resolver. La computación cuántica es ahora la mayor amenaza a las soluciones de seguridad de datos basadas en algoritmos criptográficos, como lo son los SDBs. Aunque por ahora no existe una computadora cuántica con el suficiente poder de cómputo para atacar a los sistemas de cifrado como los que se usan en un SDB, se estima que en un futuro cercano, en 2030 según expertos [8], se cuente con dicha capacidad. El impacto en la seguridad del cifrado simétrico no será tan alto como lo será para el cifrado basado en atributos y para otros tipos de cifrado que basan su seguridad en la dificultad de resolver problemas matemáticos, ya que el poder de cómputo cuántico podrá resolver dichos problemas. Por ello, los mecanismos de seguridad que se apoyan en el cifrado, como los SDB, deben fundamentar su seguridad en una criptografía postcuántica [3], misma que ya se viene desarrollando desde 2014.

Existe un ataque llamado *harvest now, decrypt later*, que en español literalmente se traduce como *colecta datos cifrados ahora, descífralos después*. Esto es, los datos cifrados, al ser ilegibles, aunque estén disponibles para un atacante, no tienen ninguna utilidad para dicho atacante. Pero el atacante podría recolectarlos ahora y, cuando se tenga ya una computadora cuántica con suficiente poder computacional, descifrar esos datos recolectados. Por ejemplo, en la Figura 9, si un atacante (como el proveedor del servicio de almacenamiento) hace un respaldo de todos los SDBs, más adelante, con la ayuda de una computadora cuántica, éste podría abrir todos esos SDBs y, por lo tanto, tener acceso a los datos que, para ese entonces, pudieran resultar de alguna utilidad. Por ejemplo, los SDBs asociados a altos mandos militares o políticos de hoy podrían revelarse en un futuro cercano (8 años aproximadamente). Cabe destacar que en el ámbito de la salud se recomienda conservar los datos médicos de una persona (i.e., historia clínica) por al menos 5 años y hasta 15 años, incluso, después de su muerte.

Por tanto, en este capítulo los autores afirmamos que uno de los principales retos para la seguridad de los datos en el sector salud es contar con mecanismos de seguridad robustos, no solo bajo los modelos de ataque actuales, sino también para aquellos modelos de ataque que se vislumbran en un futuro no muy lejano. Por otro lado, si bien los SDBs son ahora eficientes y seguros, solamente cubren los servicios de confidencialidad, integridad y control de acceso. Sin embargo, es necesario tener en cuenta que existen otros requerimientos de seguridad en salud, tales como la trazabilidad. Es muy deseable explorar el desarrollo de métodos efectivos que pudieran garantizar estos servicios, como puede ser la incorporación adecuada de tecnologías disruptivas como *Blockchain* [16].

8. Conclusiones

En este capítulo hemos introducido el concepto de sobres digitales con capacidades de búsqueda (SDB). Se trata de una abstracción que permite garantizar dos servicios de seguridad principales: la confidencialidad y el control de acceso a datos sensibles. Por ello, los SDBs son idóneos para proteger la seguridad y privacidad de datos médicos. Al estar basados en dos capas criptográficas, una fundamentada en el cifrado simétrico (rápido para cifrado) y otra en el cifrado basado en atributos (efectivo para la distribución de llaves y el control de acceso criptográfico), los SDBs integran controles de acceso de grano fino aplicables, incluso, a grandes colecciones de datos, como ocurre en el ámbito del *big data* en salud.

La fortaleza de los SDBs está probada y recae en la seguridad de ambas capas de cifrado. Su eficiencia recae en la efectividad de los patrones de paralelismo que se usan en el despliegue de los SDB, bajo la premisa de que existen recursos de cómputo disponibles (para explotar el paralelismo de datos y de tareas). La alta seguridad y eficiencia de los SDBs los hacen viables para proveer los requerimientos de confidencialidad y de control de acceso que demanda el tratamiento de datos sensibles médicos, tal como lo exige la norma mexicana NOM-024-SSA3-2012. Ante un desarrollo continuo de capacidades de un computador cuántico, los esquemas de seguridad de datos basados en cifrado se ven amenazados en el corto plazo. El trabajo futuro se está enfocando en analizar y desarrollar metodologías eficientes para incorporar cifrado postcuántico en el diseño de los SDBs.

Agradecimientos

Este trabajo forma parte del Proyecto No. 41756 CONACYT - PRONAII Ciencia de Datos en Salud *"Plataforma tecnológica para la gestión, aseguramiento, intercambio y preservación de grandes volúmenes de datos en salud y construcción de un repositorio nacional de servicios de análisis de datos de salud"*, financiado por FORDECYT-PRONACES.

Referencias

[1] Aubrey Alston. *Attribute-Based Encryption for Attribute-based Authentication, Authorization, Storage, and Transmission in Distributed Storage Systems*. Inf. téc. arXiv:1705.06002v1. Cornell University, 2017. DOI: 10.48550/arXiv.1705.06002.

[2] Elaine Barker. *Recommendation for Key Management. Part 1: General*. Inf. téc. National Institute of Standards and Technology, 2020. DOI: 10.6028/NIST.SP.800-57pt1r5.

[3] Johannes Buchmann, Kristin Lauter y Michele Mosca. "Postquantum cryptography—state of the art". En: *IEEE Security & Privacy* 15.4 (2017), págs. 12-13. DOI: 10.1109/MSP.2017.3151326.

[4] Diana Elizabeth Carrizales-Espinoza, José Luis González-Compeán y Miguel Morales-Sandoval. "Zamna: a tool for the secure and reliable storage, sharing, and usage of large data sets in data science applications". En: *2022 IEEE Mexican International Conference on Computer Science (ENC)*. IEEE, 2022. ISBN: 978-1-6654-7347-7. DOI: 10.1109/ENC56672.2022.9882938.

[5] Yunling Wang, Jianfeng Wang, Xiaofeng Chen. "Secure Searchable Encryption: A Survey". En: *Communications and Information Networks* Vol. 1. No. 4 (2016), págs. 52-65. DOI: 10.11959/j.issn.2096-1081.2016.043.

[6] Cisco Networking Academy. *Introduction to Cybersecurity*. Inf. téc. Cisco Systems, Inc., 2016.

[7] Cisco Networking Academy. *Cybersecurity Essentials*. Inf. téc. Cisco Systems, Inc., 2017.

[8] Vikas Hassija et al. "Present landscape of quantum computing". En: *IET Quantum Communication* 1.2 (2020), págs. 42-48. DOI: 10.1049/iet-qtc.2020.0027.

[9] Melissa Brigitthe Hinojosa-Cabello. "An Attribute-Based Encryption Scheme for Storage, Sharing and Retrieval of Digital Documents in the Cloud". Tesis de mtría. Cinvestav, 2020.

[10] Richard Kuhn et al. *Introduction to Public Key Technology and the Federal PKI Infrastructure*. Inf. téc. National Institute of Standards and Technology, 2001.

[11] Badrinarayanan Lakshmiraghavan. *Pro ASP.NET Web API Security. Securing ASP.NET Web API*. Ed. por Apress Media, LLC. 1.ª ed. Springer, 2013. 416 págs. ISBN: 978-1-4302-5782-0. DOI: 10.1007/978-1-4302-5783-7.

[12] Elaine Barker, William Barker, Annabelle Lee. *Guideline for Implementing Cryptography in the Federal Government*. Inf. téc. National Institute of Standards and Technology, 2005.

[13] Antonis Michalas. "The Lord of the Shares: Combining Attribute-Based Encryption and Searchable Encryption for Flexible Data Sharing". En: *SAC '19: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. Association for Computing Machinery, 2018. ISBN: 978-1-4503-5933-7. DOI: 10.1145/3297280.3297297.

[14] Praveen Kumar Premkamal, Syam Kumar Pasupuleti y Pja Alphonse. "Attribute Based Encryption in Cloud Computing: A Survey, Gap Analysis, and Future Directions". En: *Network and Computer Applications* 108 (2018), págs. 37-52. DOI: 10.1016/j.jnca.2018.02.009.

[15] Conacyt PRONACES. *Plataforma tecnológica para la gestión, aseguramiento, intercambio y preservación de grandes volúmenes de datos en salud y construcción de un repositorio nacional de servicios de análisis de datos de salud. PRONACES Salud, FORDECYT 2019-06 CONACyT, proyecto número 41756.* http://adaptivez.org.mx/e-SaludData/. 2022.

[16] Nabil Rifi et al. "Towards using blockchain technology for eHealth data access management". En: *2017 fourth international conference on advances in biomedical engineering (ICABME)*. IEEE, 2017. ISBN: 978-1-5386-1642-0. DOI: 10.1109/ICABME.2017.8167555.

[17] Amit Sahai y Brent Waters. "Fuzzy Identity-Based Encryption". En: *Advances in Cryptology – EUROCRYPT 2005*. Springer Berlin Heidelberg, 2005. ISBN: 978-3-540-32055-5. DOI: 10.1007/11426639_27.

[18] Douglas Selent. "Advanced encryption standard". En: *Rivier Academic Journal* 6.2 (2010), págs. 1-14.

[19] Víctor Jesús Sosa-Sosa et al. "Protecting Data in the Cloud: An Assessment of Practical Digital Envelopes from Attribute based Encryption". En: *KDCloudApps 2017*. SciTePress, 2017. ISBN: 978-989-758-255-4. DOI: 10.5220/0006484603820390.

[20] Hui Bin Yin et al. "CP-ABSE: A Ciphertext-Policy Attribute-Based Searchable Encryption Scheme". En: *IEEE Access* 7 (2019), págs. 5682-5694. DOI: 10.1109/ACCESS.2018.2889754.

[21] Rui Zhang, Rui Xue y Ling Liu. "Searchable encryption for healthcare clouds: a survey". En: *IEEE Transactions on Services Computing* 11.6 (2017), págs. 978-996. DOI: 10.1109/TSC.2017.2762296.

[22] Eduardo Palma Ávila. "Criptografía Basada en Hardware". En: *Revista Seguridad. Cultura de prevención para TI*. No. 21. Universidad Nacional Autónoma de México, jun. de 2014.