



UNIVERSIDAD POLITÉCNICA DE MADRID

MÁSTER UNIVERSITARIO EN
CIBERSEGURIDAD 2020

TRABAJO FIN DE MÁSTER:

Diseño y evaluación de
procedimientos para informática
forense sobre bases de datos

Eduardo Rodríguez Hernández

TRABAJO FIN DE MÁSTER

TÍTULO: Diseño y evaluación de procedimientos para informática forense sobre bases de datos.

AUTOR: D. Eduardo Rodríguez Hernández

TUTOR 1/DIRECTOR: D. Borja Bordel Sánchez

Escuela Técnica Superior de Ingenieros de Telecomunicación

Escuela Técnica Superior de Ingeniería y Sistemas de Telecomunicación

Escuela Técnica Superior de Ingenieros Informáticos

Escuela Técnica Superior de Ingeniería de Sistemas Informáticos

TRIBUNAL:

PRESIDENTE:

VOCAL:

SECRETARIO:

FECHA DE LECTURA:

CALIFICACIÓN:

Fdo: El Secretario del Tribunal

Agradecimientos

A mi familia por haberme apoyado siempre y haber estado ahí cuando los necesitaba.

A mis compañeros y amigos, especialmente a los nuevos, por haberme ayudado y acompañado a lo largo de este duro y disparatado año de máster 2019/2020.

A los profesores que han ayudado a mi desarrollo en el campo de estudio de este trabajo y en especial a mi tutor por haberme guiado e incentivado a lo largo del mismo.

Resumen

Los sistemas de bases de datos ganan cada vez más importancia en el mundo de la informática, con el crecimiento de los servicios web y aplicaciones que cada vez más y más se implementan en todos y cada uno de los campos cotidianos, desde sistemas de gran importancia como los de gestión del gobierno central de países hasta minucias como los sistemas de las aplicaciones más sencillas.

En caso de encontrar un sistema de estas características comprometido por algún tipo de malfuncionamiento, corrupción o malware, en muchos casos es de vital importancia el conocimiento de la procedencia del fallo. Para esclarecer los hechos de un problema como el mencionado, se hace uso de procedimientos forenses sobre los sistemas, en el caso de los sistemas de bases de datos encontramos diversos procesos especialmente dependiendo del fabricante.

El objetivo de este trabajo es diseñar, aplicar y evaluar una serie de procedimientos para informática forense sobre un sistema de bases de datos libre, la investigación se realiza con la intención de ofrecer una catálogo de procedimientos que ayuden a la recolección y análisis de evidencias forenses generadas por una base de datos en un caso lo más fiel a la realidad posible dentro de un entorno de laboratorio.

Abstract

Database systems are becoming increasingly important in the world of computing, with the growth of web services and applications that are increasingly being implemented in each and every field of daily life, from major systems such as central government management of countries, to minor systems such as the simplest of applications.

In the event of finding a system of these characteristics compromised by some type of malfunction, corruption, or malware, in many cases it is of vital importance to discover the origin of the failure. In order to clarify the causes of a problem with this structure, forensic procedures are used on the systems. In the case of database systems, several processes are found, especially depending on the manufacturer of the system.

The objective of this project is to design, apply and evaluate a series of procedures for computer forensics on a free database system. The investigation is carried out with the intention of offering a catalogue of procedures that help the collection and analysis of forensic evidence generated by a database in a case as faithful to reality as possible within a laboratory environment.

Contenido

Agradecimientos	i
Resumen	iii
Abstract	v
Contenido	vii
Listado de figuras	ix
Listado de tablas.....	xi
Acrónimos.....	13
1 Introducción	14
1.1 Contexto y justificación	14
1.2 Objetivos	15
1.3 Estructura	16
1.4 Competencias cubiertas	17
2 Bases de datos	19
2.1 Modelos y Clasificaciones.....	19
2.2 Lenguaje SQL.....	22
2.3 Seguridad en Bases de datos.....	25
2.4 Sistemas de gestión de bases de datos.....	27
3 Informática forense	33
3.1 ¿Qué es la informática forense?	34

3.2	Evidencias forenses.....	36
3.3	Modelos de informática forense.....	37
3.4	Estándares Nacionales.....	40
4	Desarrollo de una metodología de análisis forense.....	44
4.1	Características y requisitos.....	44
4.2	Introducción.....	45
4.3	Fase 1: Conservación del entorno.....	46
4.4	Fase 2: Identificación y recolección de evidencias.....	47
4.5	Fase 3: Conservación de las evidencias.....	48
4.6	Fase 4: Análisis de las pruebas.....	49
4.7	Fase 5: Evaluación de las pruebas.....	51
4.8	Fase 6: Realización del informe.....	52
5	Caso Práctico.....	54
5.1	Caso práctico.....	54
5.2	Preparación del caso práctico.....	55
5.3	Aplicación de la metodología.....	56
5.4	Informe Forense del caso.....	57
6	Conclusiones y líneas futuras.....	61
6.1	Repaso y conclusiones.....	61
6.2	Líneas futuras de trabajo.....	62
7	Anexos.....	63
7.1	Anexo 1: capturas del servicio web.....	63
7.2	Anexo 2: mensajes de error, diagrama de la base de datos del caso, descripción del servicio.....	65
7.3	Anexo 3: Capturas de pantalla.....	67
8	Bibliografía.....	69

Listado de figuras

Figure 1: Ejemplo de base de datos relacional.....	20
Figure 2: Logo MySQL.....	29
Figure 3: Logo SQLite.....	30
Figure 4: Logo MongoDB.....	31
Figure 5: Logo Redis.....	32
Figure 6: Diagrama del principio de Locard.....	34
Figure 7: Modelo DFRW.	37
Figure 8: Diagrama del modelo de análisis forense militar.	38
Figure 9: Diagrama Casey 2004.	39
Figure 10: Fases de la metodología.....	45
Figure 11: Sistema de roles.	63
Figure 12: Obra de un artista.	64
Figure 13: ejemplo de propuesta.	64
Figure 14: Ejemplo de mensajería.	64
Figure 15: Error en la base de datos.....	65
Figure 16: Diagrama de la base de datos.	65
Figure 17: Stellar Repair for SQLite.	67
Figure 18: SysTools SQLite Database Recovery.	67
Figure 19: Binwalk sobre base de datos.	68

Figure 20: entropía del fichero de bases de datos.	68
---	----

Listado de tablas

Tabla 2.2.2.1: operaciones DML	24
Tabla 3.4.1.1: Sub-normas UNE 71505:2013.	41
Tabla 3.4.3.1: Algunas herramientas para realizar análisis sobre bases de datos.	50
Tabla 5.1.2.1: cadena de custodia de evidencias.....	58

Acrónimos

Si no hay lista de acrónimos, se debe quitar esta página. Ejemplo de acrónimos:

TFM	Trabajo Fin de Máster
UPM	Universidad Politécnica de Madrid
SQL	Structured Query Language
NoSQL	Sistema no solo SQL
DML	Data Manipulation Language
DDL	Data Definition Language
DCL	Data Control Language
SGBD	Sistema de gestión de Base de datos
SSL	Secure Socket Layer
CMS	Content Management System
ACID	Atomicidad, Consistencia, Aislamiento, Durabilidad
DFRW	Digital Forensic Research Workshop
AENOR	Asociación Española de Normalización
UNE	Una Norma Española
RFC	Request For Comments
DBMS	Database Management System
SSH	Secure Shell

1

Introducción

1.1 Contexto y justificación

Este trabajo está desarrollado para el máster en ciberseguridad de la universidad politécnica de Madrid, se funda en el campo de la ingeniería forense incluyendo recolección y análisis de evidencias digitales, especialmente orientado hacia los sistemas de bases de datos.

Actualmente los procesos de recolección y análisis forense orientados hacia las bases de datos no se encuentran estandarizados, aun así, es posible encontrar referencias a ciertas pautas y metodologías a seguir recomendadas por diversas organizaciones tanto nacionales como internacionales e incluso los proveedores de sistemas de las propias bases de datos, que pueden ayudar en el caso de ser necesario llevar a cabo una investigación de estas características. Aun existiendo este tipo de pautas y metodologías a seguir y estar recomendados por organizaciones de renombre, no existe certeza ante aspectos legales de que el seguimiento de estas pautas y metodologías garanticen la completitud de la investigación.

Este trabajo por tanto se basa en intentar probar un nuevo acercamiento a una metodología orientada específicamente a las investigaciones forenses sobre sistemas de bases de datos, unificando procedimientos, prácticas y orientaciones en un solo documento y probando en un caso práctico desarrollado en un entorno de laboratorio, que permita obtener una serie de conclusiones y resultados sobre las metodologías aplicadas.

1.2 Objetivos

En el campo de la ingeniería forense existe mucha diversidad con respecto a las evidencias encontradas en cada caso particularmente, puesto que cada compañía u organización posee sistemas distintos y administrados de distinta manera, es por ello que, en función de valores como el sistema operativo de la máquina afectada, sistema de gestión de bases de datos, etc. Puede verse variado el proceso de acercamiento hacia las evidencias.

El objetivo último del proyecto es dar y probar una metodología para el análisis forense de sistemas de bases de datos en un entorno de laboratorio lo más cercano a un caso real posible, para poder orientar a posibles investigadores en el evento de necesitar realizar una investigación hacia uno de estos sistemas y hacerlo de la forma más genérica posible teniendo en cuenta los estándares actuales. Este objetivo se podría desglosar en los siguientes objetivos más específicos:

1. Realizar un estudio sobre los actuales sistemas de bases de datos.
2. Investigar acerca de las vigentes técnicas de Informática forense.
3. Desarrollo de una metodología de análisis forense teniendo en cuenta los anteriores puntos
4. Prueba de la anterior metodología de análisis forense contra un caso práctico preparado.
5. Hallar conclusiones sobre la metodología y su prueba.

1.3 Estructura

Este trabajo de fin de máster comprende seis capítulos contando el capítulo de la introducción y el capítulo final para anexos necesarios.

- En el segundo capítulo de esta memoria se detalla información al respecto de los sistemas de bases de datos en general, su funcionamiento habitual, lenguajes usados para controlarlas y sistemas de seguridad que se poseen por defecto para protegerlos, es necesario debido a la orientación de este proyecto debido a que se debe tener una base de conocimiento previo antes de aplicar conocimiento de ingeniería forense sobre un sistema de bases de datos.
- En el tercer capítulo se habla del estado de la ingeniería forense en la actualidad, incluyendo algunos modelos/metodologías habituales y de referencia, tratamiento de evidencias, estándares nacionales de utilidad.
- Durante el cuarto capítulo se expone la metodología a seguir en el caso práctico de este proyecto.
- En el quinto capítulo se define el caso de laboratorio creado, con la metodología seguida y el informe de la investigación principal después de la puesta en situación.
- En el sexto capítulo se relatan las conclusiones tanto del informe de la investigación como de la aplicación de la metodología y las posibles líneas futuras de trabajo.

1.4 Competencias cubiertas

Como competencias cubiertas específicamente por este trabajo de fin de máster encontramos (definidas en el plan de estudio de la titulación):

- Básicas y generales:
 - “CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.”
 - “CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.”
 - “CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.”
 - “CG2 - Dotar al alumno del conocimiento de los distintos tipos de amenazas que pueden afectar a una organización y sus consecuencias en diferentes escenarios sociales, económicos e industriales y dotarle de la capacidad de aplicar las técnicas de análisis y gestión de todo tipo de riesgos para definir e implantar las salvaguardas necesarias para mitigar o eliminar sus impactos hacer resiliente a la organización.”
- Transversales:
 - “CT4 - Organización y planificación.”
 - “CT5 - Gestión de la información.”
 - “CT9 - Capacidad de análisis y síntesis.”
 - “CT10 - Resolución de problemas.”
 - “CT12 - Aprendizaje autónomo, adaptación a nuevas situaciones y motivación por el desarrollo profesional permanente.”
- Específicas:

- “CE2 - Capacidad para diseñar estrategias, políticas y normativas de ciberseguridad corporativa.”
- “CE6 - Capacidad de aplicar las principales metodologías y técnicas de seguridad en el desarrollo del software y sistemas informáticos.”
- “CE8 - Capacidad para abordar técnicamente la gestión de un incidente de ciberseguridad, incluyendo análisis de malware, análisis forense e ingeniería inversa.”
- “CETFM - Capacidad de realizar un trabajo o proyecto individual integrando y relacionando las competencias adquiridas en las distintas asignaturas del máster, junto con la capacidad de defenderlo en público ante un grupo de personas expertas en el tema del trabajo.”

2 Bases de datos

Una base de datos representa un conjunto de datos que por lo general poseen un contexto común y que se almacenan para ser usados en otro momento, ya sea por aplicaciones o por consultas directas a la base de datos.

Las bases de datos se clasifican siguiendo una serie de criterios, pero por norma general existen tres criterios principales que son: su forma o modelo de administrar los datos, su variabilidad y su contenido.

2.1 Modelos y Clasificaciones

Cuando se habla de un **modelo de administración de datos** de una base de datos, básicamente se está haciendo referencia a la forma de almacenar los datos para su posterior recuperación por parte de la base de datos, existen diversos modelos como, por ejemplo:

- **Relacionales:**
 - el modelo relacional de bases de datos es el modelo de referencia actualmente y el más usado debido a su capacidad para representar con facilidad modelos reales de problemas y su facilidad de administración.
 - Este paradigma se basa en hacer uso de “relaciones” o lo que sería conjuntos de datos relacionados unos con otros a los cuales se denomina “tuplas”.
 - Una de las grandes ventajas de este modelo es que no se debe tener en cuenta la forma ni el lugar en el que se almacenen los datos, lo cual hace que estos datos sean más fáciles e intuitivos de usar.
 - El lenguaje de uso habitual de este modelo de bases de datos es SQL

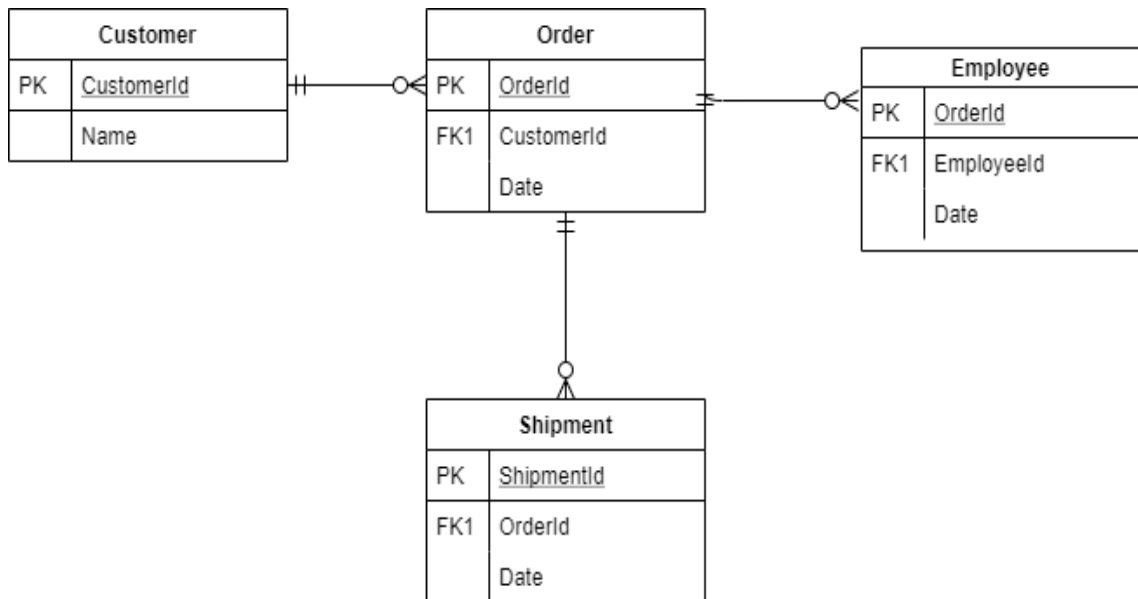


Figure 1: Ejemplo de base de datos relacional.

- **Transaccionales:** Las bases de datos transaccionales son bases de datos las cuales están orientadas a operaciones a alta velocidad. Tienen la ventaja de que debido a su orientación envían y recuperan datos a velocidades óptimas, pero al ser esa su principal característica, se descuidan los aspectos de redundancia y duplicación de datos (a pesar de que no deban ser de importancia en sistemas que usan estas bases de datos.)
- **Jerárquicas:** Las bases de datos jerárquicas se estructuran como un árbol invertido, es decir, existe un nodo padre que almacena información y puede tener varios hijos que a su vez tengan otros hijos, a los nodos finales o nodos sin hijos se les considera “hojas” del árbol. Este modelo proporciona utilidad para grandes volúmenes de información y datos compartidos, pero tiene grandes problemas con la redundancia de datos, debido a que los datos se pueden repetir en otro nivel y rama del árbol sin identificación por parte del nodo redundante.

- **Orientadas a objetos:** Aunque relativamente nuevo, este modelo de bases de datos está ganando popularidad poco a poco, se basa en la premisa de almacenar objetos completos en la base de datos incluyendo su estado y comportamiento, esto lo convierte en un modelo bastante atractivo para los paradigmas de la programación actual, además incluye conceptos importantes para estos paradigmas de programación orientada a objetos, como son:
 - **Encapsulamiento** para evitar conflictos y accesos incorrectos a los datos proporcionándoles independencia entre sí y ocultación.
 - **Herencia** para garantizar los comportamientos y características que los datos reciben de clases jerárquicas.
 - **Polimorfismo** para poder garantizar que los objetos mantienen la aplicabilidad a diferentes operaciones y las operaciones a los diferentes tipos de objetos que se puedan encontrar.

Se admite en este tipo de bases de datos la creación de funciones específicas para la realización de operaciones personalizables por parte de los usuarios o administradores de la base de datos en cuestión, aunque se debe especificar separadamente.

Según su **variabilidad** se encuentran bases de datos **estáticas** (cuyos datos no se modifican, es decir, están preparadas solo para operaciones de lectura) y **dinámicas** (cuyos datos si se pueden modificar, es decir, están preparadas para operaciones de escritura, actualización y borrado además de lectura).

Según su **contenido** aparecen 4 clasificaciones para las bases de datos, que son:

- **Bibliográficas:** usadas para contener datos meramente aditivos frente a otros datos contenidos en otras bases de datos.
- **De texto completo:** se usan para contener datos en su completitud.
- **Directorios:** contienen datos de uso personal o empresarial, un gran ejemplo son las guías telefónicas, que contienen datos de personas y compañías.
- **De información química o biológica:** contienen datos e información de tipo biológico/médico o químico.

2.2 Lenguaje SQL

SQL o 'Structured Query Language' es un lenguaje de programación informático, especialmente diseñado para realizar la gestión sobre bases de datos. La mayoría de las bases de datos hacen uso de este lenguaje de gestión para su administración, pero especialmente las relacionales que son las que se verán más fuertemente en el ámbito de este trabajo.

La principal característica de este lenguaje es que hace uso de álgebra relacional para generar consultas usadas para gestionar los datos dentro de una base de datos.

SQL en sí consiste en tres lenguajes para llevar a cabo todas las operaciones, un lenguaje de definición de datos, un lenguaje de manejo de datos y un lenguaje de control de datos.

2.2.1 Data Definition Language

Un **Lenguaje de definición de datos (DDL, Data Definition Language)** para poder llevar a cabo la definición de las estructuras que van a almacenar los datos y dar definición también a los procedimientos que permitan consultar y modificar los datos contenidos en dichas estructuras.

El lenguaje de definición de datos de SQL hace uso de una serie de sentencias que sirven para realizar las operaciones básicas de creación, borrado y modificación, dichas sentencias se basan en los siguientes comandos:

- **CREATE:** se usa para crear nuevas bases de datos, tablas, índices o procedimientos almacenados. Las sentencias tienen la siguiente estructura:
 - *CREATE 'tipo de objeto' 'nombre del objeto'*
- **DROP:** se usa para eliminar estructuras de datos o elementos de la base de datos de manera sencilla. Las sentencias tienen la siguiente estructura:
 - *DROP 'tipo de objeto' 'nombre del objeto'*
- **ALTER:** se usa principalmente para modificar aspectos de las estructuras de datos o elementos de la base de datos, aunque también puede ser usada para crear y eliminar estructuras o elementos nuevos. Las sentencias siguen el orden siguiente:
 - *ALTER 'tipo de objeto' 'operación' 'dato a añadir/eliminar/modificar' 'nombre del dato'*

Adicionalmente se encuentra una sentencia digna de comentar, se habla del caso de la operación '**TRUNCATE**', es una operación exclusiva de las tablas y sirve para eliminar el contenido dentro de una tabla de la base de datos. Su existencia se debe a que es una operación de limpieza, por tanto, es mucho más rápida que la operación **DROP** o **DELETE** (se verá en el siguiente punto), pero no admite ningún tipo de cláusula para especificar, por tanto, solo sirve para eliminar tablas en su completitud. Se la considera una operación DDL y no una DML puesto que el funcionamiento del comando se basa en borrar la tabla y volver a crearla, pero vacía, de forma que no se ejecuta ninguna transacción.

2.2.2 Data Management Language

Un **Lenguaje de manejo de datos (DML, Data Management Language)** para permitir a los usuarios o programadores llevar a cabo las consultas pertinentes referentes a modificación, creación, etc. De datos dentro la base de datos.

Este lenguaje se basa en cuatro sentencias básicas, que son **SELECT**, **INSERT**, **DELETE** y **UPDATE**.

Las consultas de tipo **SELECT** permite recuperar datos de la base de datos, lo más habitual es que se recuperen de una de las tablas contenida en la base de datos. **SELECT** como consulta permite el uso de ciertas palabras clave que se usan para acotar las búsquedas, de forma que sean más específicas y menos costosas, algunas de esas palabras (las más usadas habitualmente) corresponden con las de la tabla 1 vista a continuación.

Tabla 2.2.2.1: operaciones DML

Palabra	Acción
WHERE	Hace uso de los operadores lógicos AND y OR para añadir condiciones que deben cumplirse a la sentencia SELECT.
FROM	De uso necesario al aplicar un WHERE en las sentencias, indica de donde se deben sacar los datos a los que la consulta hace referencia
ORDER BY	Se usa en conjunción con otras palabras de acotación de búsqueda para ordenar el resultado de la cuestión de forma ascendente o descendente en función de un campo de la base de datos.

Las consultas de tipo **INSERT** son sentencias de uso básico para realizar inserciones de valores en las tablas de las bases de datos. Se deben dar como valores todos los campos que sean obligatorios según la base de datos.

Las consultas de tipo **DELETE** es una sentencia que se usa para eliminar registros pertenecientes a una tabla indicada. Se puede hacer uso de cláusulas “**WHERE**” para establecer filtros y eliminar registros específicos de la tabla.

Las consultas que hagan uso de la sentencia **UPDATE** se utilizan para modificar registros actuales de la tabla, para eso se hace uso de la palabra auxiliar **SET** y se añade como parámetro una clave de registro y el nuevo valor por el que se va a actualizar. En caso de no encontrar registros que coincidan con la clave proporcionada en el parámetro ningún valor de la tabla se ve afectado.

2.2.3 Data Control Language

Un **Lenguaje de control de datos (DCL, Data Control Language)** es un componente de SQL usado para permitir que el administrador del sistema pueda tener control sobre los datos y las consultas realizadas sobre ellos.

Los comandos básicos de este lenguaje son **GRANT** y **REVOKE**, los cuales se usan para dar o quitar permisos a los usuarios sobre las actividades y accesibilidad en la base de datos. En especial afectan a los apartados vistos en DML y DDL, incluyendo sentencias como **INSERT**, **SELECT**, **DELETE**, etc.

2.3 Seguridad en Bases de datos

En el contexto de las bases de datos, que son el objeto de estudio de este proyecto, debemos tener en cuenta que existen 3 tipos de medidas de seguridad, que son: **físicas**, **personales** y **propias del sistema de gestión de bases de datos**. En este caso de estudio las medidas que realmente interesan son las últimas, es decir, las medidas de seguridad que se aplican en el sistema de gestión de bases de datos por norma general.

Para los sistemas de gestión de bases de datos encontramos dos tipos principales de seguridad, que son **direccional** y **obligatoria**.

- **Direccional**: es el tipo de seguridad usada para dar y quitar privilegios a los posibles usuarios, ya sea a nivel de los archivos de las bases de datos, los registros o los campos específicos tanto en modo consulta como modificación.
- **Obligatoria**: se usan para conseguir clasificar los datos y usuarios en diversos niveles de seguridad que se usarán para garantizar la igualdad de los niveles y posteriormente aplicar una política de seguridad con mecanismos de protección por niveles.

Una vez se cumplan estos tipos de seguridad, deben estar garantizados una serie de requisitos, como son garantizar la capacidad de reconstrucción de los datos de la base de datos, susceptibilidad a ser pasados por procesos de auditoría standard, no debe ser posible la intrusión a los sistemas (al menos de manera sencilla), se debe comprobar la autorización de las acciones antes de su realización y las acciones realizadas dentro del SGBD deben ser controladas para evitar acciones mal intencionadas.

2.3.1 Riesgos de seguridad para bases de datos

Las bases de datos se ven afectadas normalmente por una amplia colección de riesgos, muchos de estos riesgos se ven minimizados por sistemas de seguridad de red comunes, como son los cortafuegos y sistemas de detección de intrusos. Por otro lado, encontramos posibles riesgos que afectan a sistemas de bases de datos intrínsecamente como son:

- Infecciones de malware o ataques intencionados que puedan generar corrupción, eliminación o adición de registros de la base de datos, cifrado de los sistemas de bases de datos, etc.
- Acciones no autorizadas o mal intencionadas.
- Fallos en el diseño que puedan conllevar sobrecargas, errores en el sistema, etc.
- Vulnerabilidades de seguridad en el sistema o programas asociados que puedan generar accesos indebidos o no autorizados a la propia base de datos.

2.3.2 Procedimientos habituales.

Por norma general cada compañía es libre de desarrollar y utilizar su propia política de protección de bases de datos, de forma que una actúa de forma distinta a la hora de proceder a proteger sus sistemas de bases de datos. Aun así, se debe estar seguro de que se cumplen una serie de requisitos mínimos que vienen detallados por legislaciones, especialmente teniendo en cuenta el apartado de la privacidad de los usuarios que hagan uso de los servicios de la compañía. Lo habitual es que se sigan las recomendaciones de seguridad proporcionadas por el fabricante o proveedor de sistemas de gestión de bases de datos, además se debe tener una fuerte política de gestión del sistema contando con administradores, usuarios fuertemente diferenciados mediante roles, sistemas de registro de actividades, etc. Los puntos más generales a la hora de proteger un sistema de bases de datos son por norma general:

- Copias de seguridad para mantener los datos lo más actualizados posibles en caso de pérdida irrecuperable.
- Encriptación, tanto para los registros como para los sistemas de la base de datos donde pueda aplicarse.
- Autenticación para poder controlar los usuarios que acceden y los permisos que tienen sobre la información guardada en la base de datos.
- Integridad con sus correspondientes controles para verificar que la información guardada es correcta y está completa.

2.4 Sistemas de gestión de bases de datos.

Los sistemas de gestión de bases de datos o SGBD representan un conjunto de programas que permiten realizar una serie de funciones como son el almacenamiento, modificación y extracción de registros de una base de datos.

Poseen varias ventajas como son por ejemplo la simplificación de la programación de sistemas que usen SGBD, organización de los datos de la base de datos, se puede hacer uso de políticas de control para los errores y cambios en la base de datos y además, suelen proveer lenguajes de consulta e interfaces de sencillo uso para los usuarios. Pero por otro lado pueden llegar a ser excesivamente complejos para sistemas muy sencillos, pueden llegar a ser difíciles de utilizar y administrar, por tanto, los usuarios deben tener conocimientos básicos sobre manejo de SGBDs y suelen requerir de hardware adicional que conlleva un coste.

Estos sistemas por norma general garantizan completamente la atomicidad, consistencia, aislamiento y durabilidad (principios ACID) dentro de la base de datos, además de permitir administración de acceso de los usuarios y permitir recuperación de información en caso de haber cualquier problema en el sistema que derive en corrupción de datos o eliminación errónea de registros.

Los SGBD poseen una serie de objetivos que deben ser cumplidos que son:

- Los datos de la base de datos deben ser totalmente independientes del diseño, de forma que cualquier modificación en el diseño de esta no conlleve cambios en las aplicaciones que hacen uso de la base de datos.
- Los SGBD deben poseer un control de las transacciones que se realizan en la base de datos, de forma que, si se produce cualquier error en una cuestión, exista una forma de lidiar con ella sin perjudicar a la base de datos.
- Los sistemas deben estar abstraídos con respecto a la información de la base de datos, es decir, deben facilitar al usuario la realización de operaciones sobre la base de datos, para ello se debe abstraer al usuario de forma que este opere siempre de la misma forma sin importancia sobre el diseño físico de la base de datos.
- Debe haber consistencia en los datos, es decir, en los casos en los que existan registros redundantes por necesidad o diseño, estos se deben actualizar correctamente.
- Los tiempos de respuesta entre el SGBD y la base de datos deben ser aceptables para los sistemas que vayan a hacer uso del sistema para realizar sus transacciones.

Los SGBD pueden clasificar dependiendo de la forma en la que administran sus datos entre Relacionales y no relacionales.

2.4.1 SGBDs Relacionales

Los SGBDs relacionales se basan en el modelo relacional de bases de datos explicado anteriormente en este documento, el cual es el modelo más utilizado en la actualidad para administrar bases de datos, existe un gran número de SGBDs que se corresponden con este modelo tanto de forma comercial como libre.

Los SGBDs más importantes para el desarrollo de este proyecto son los que posean licencia general o sean gratuitos de uso, puesto que son los más relevantes para el posterior caso práctico, en ese caso los SGBD más importantes son:

MySQL: es el SGBD de bases de datos relacionales más usado a nivel comercial por grandes compañías, se ofrece bajo licencia tanto pública como privada o comercial por parte de Oracle, lo cual permite integrarlo en ciertos sistemas de producción privados.

Sus principales ventajas residen en que es un sistema multiusuario y multihilo, lo cual lo dota de concurrencia a la hora de realizar peticiones a la base de datos, también es un sistema muy rápido en velocidades de lectura, posee un gran rendimiento y es sencillo de usar, además ofrece soporte para SSL y multiplataformas.

Sus principales desventajas se basan en la eficiencia de trabajo para bases de datos exageradamente grandes, por tanto, su escalabilidad y, por otro lado, el hecho de que, para sistemas de pequeñas empresas o sistemas de diseño muy sencillo, puede ser más costoso de implementar y mantener que otros sistemas más sencillos.



Figure 2: Logo MySQL.

SQLite: es un SGBD basado en una biblioteca escrita en C, es muy ampliamente utilizado, especialmente por aplicaciones sencillas que no requieran excesiva complejidad o cuyo diseño de la base de datos no sea muy complejo. Permite que se realicen transacciones evitando el montaje de un servidor de base de datos con sus correspondientes configuraciones.

SQLite se usa en multitud de aplicaciones, especialmente CMS y aplicaciones en lenguajes de programación como PHP debido a su optima integración y a su distribución mediante licencia de código abierto.

Las grandes ventajas de SQLite residen en su eficiencia a la hora de realizar consultas, su tamaño, puesto que, al estar basado en una biblioteca no posee el mismo tamaño que otros SGBD de implementación completa, por otro lado, es un sistema de alta portabilidad y rendimiento debido a su ligereza y compatibilidades.

SQLite soporta todos los criterios de atomicidad, consistencia, aislamiento y durabilidad de un sistema de gestión de bases de datos y además debido a su ligero peso de instalación y funcionamiento se convierte en perfecto para sistemas integrados y pequeñas empresas con aplicaciones que no requieran excesivos recursos.



Figure 3: Logo SQLite.

2.4.2 SGBDs No relacionales

Los SGBDs no relacionales son sistemas que hacen uso de bases de datos no relacionales, es decir, bases de datos que no hacen uso de estructuras preestablecidas o fijas como son las tablas, se denominan sistemas no relacionales o “NoSQL” debido a que son capaces de funcionar con lenguajes de consulta SQL pero que no son requeridos.

Se usan por norma general en entornos donde la base de datos debe estar siempre disponible, operativa y lista para manejar unos sets de datos que suelen ser muy grandes en tamaño. Las bases de datos NoSQL están muy optimizadas para operaciones de lectura e inserción, además poseen una gran escalabilidad y eficiencia en modelos de datos específicos, aunque se pierde mucha flexibilidad con respecto al tiempo de ejecución de las consultas comparando con sistemas más habituales SQL.

Los SGBD más importantes dentro de los no relacionales son:

MongoDB: Es probablemente el SGBD no relacional más usado en la actualidad, posee gran popularidad y aceptación de uso entre las tecnologías. Su sistema de base de datos distribuida y basada en ficheros es de muy sencillo uso y ampliamente apto para las aplicaciones más modernas.

MongoDB posee entre sus características más importantes la capacidad para trabajar en la nube, lo cual lo posiciona como un favorito de uso entre desarrolladores en la actualidad, además ofrece un nivel de productividad elevado.

Su mayor ventaja frente a la competición es el uso de ficheros JSON (formato de texto para intercambio de datos), lo cual le dota una gran integración con la nueva generación de aplicaciones web, además ofrece balanceo de cargas, alta escalabilidad horizontal y su codificación abierta, pero por otro lado su gran desventaja es que para aplicaciones complicadas no posee capacidad de realizar consultas lo suficientemente complejas.



Figure 4: Logo MongoDB.

Redis: Actúa de motor de base de datos, más que de SGBD, se basa en un sistema de almacenamiento de clave-valor, donde se relaciona una clave con un contenido almacenado dentro de un índice.

Su principal ventaja como sistema reside en que al contrario que otros sistemas como MongoDB, su contenido no se reduce al tipo primitivo 'string', sino que soporta tipos de datos más complejos como listas y sets de string o hashes. Otra gran ventaja de los sistemas de Redis es que hace uso de un modelo de maestro-esclavo para llevar a cabo replicación, de forma que se puede distribuir la base de datos para disminuir tiempos de carga y de transacción.



Figure 5: Logo Redis.

3

Informática forense

El objetivo de este capítulo es llevar a cabo un análisis de la informática forense actual desde un punto de vista general, para poder luego generar las bases que llevan a la creación de una metodología para el análisis forense orientándola específicamente hacia un sistema de bases de datos, especialmente teniendo en cuenta los puntos de una metodología forense genérica, pero aplicados a las estructuras actuales de sistemas de bases de datos y sus posibles configuraciones.

En este capítulo también se tratarán temas como la recolección y gestión de evidencias en un caso de investigación forense a nivel genérico y específico para las competencias del trabajo y se tratarán algunos modelos de ingeniería forense, los cuales ayudarán al posterior desarrollo de la metodología a utilizar en el caso práctico de este trabajo.

Por otro lado, se deben tratar algunos estándares nacionales e internacionales que se pueden seguir para realizar un análisis forense de forma satisfactoria, puesto que es necesario tener una serie de conocimientos sobre normativas y formas de operar al respecto antes de lanzarse a hacer una investigación de estas características.

3.1 ¿Qué es la informática forense?

La informática forense es un área relativamente reciente pero que, a pesar de esto, se encuentra en rápido crecimiento, especialmente teniendo en cuenta el exponencial aumento de la ciberdelincuencia experimentado en los últimos años.

La informática forense es un área que requiere amplios conocimientos por parte de los investigadores, puesto que no solo se limita al ámbito técnico hablando de sistemas informáticos, sino también a otros ámbitos como son el legal y el criminal.

Para poder llevar a cabo correctamente una investigación forense en el campo de la informática, se deben tener claros los objetivos, los cuales se corresponden con identificar, adquirir, recuperar/conservar evidencias digitales y presentar informes con hechos objetivos deducidos a partir del análisis de las previas evidencias. Una investigación de informática forense siempre va a ser una investigación muy dependiente de diversos factores, desde el equipo/red sobre el que se realice hasta el momento de la línea temporal del daño que se ha encontrado y se intenta investigar.

La disciplina de la ingeniería forense se basa en el principio del intercambio de **Locard**, el cual indica que cualquier objeto que haya estado implicado en una escena de un crimen, deja un rastro ya sea en la escena del crimen o en la víctima con la que ha estado en contacto, el objetivo en el caso de la informática forense es obtener evidencias dentro de esa escena del crimen que permitan averiguar la sucesión de hechos en dicha escena del crimen.

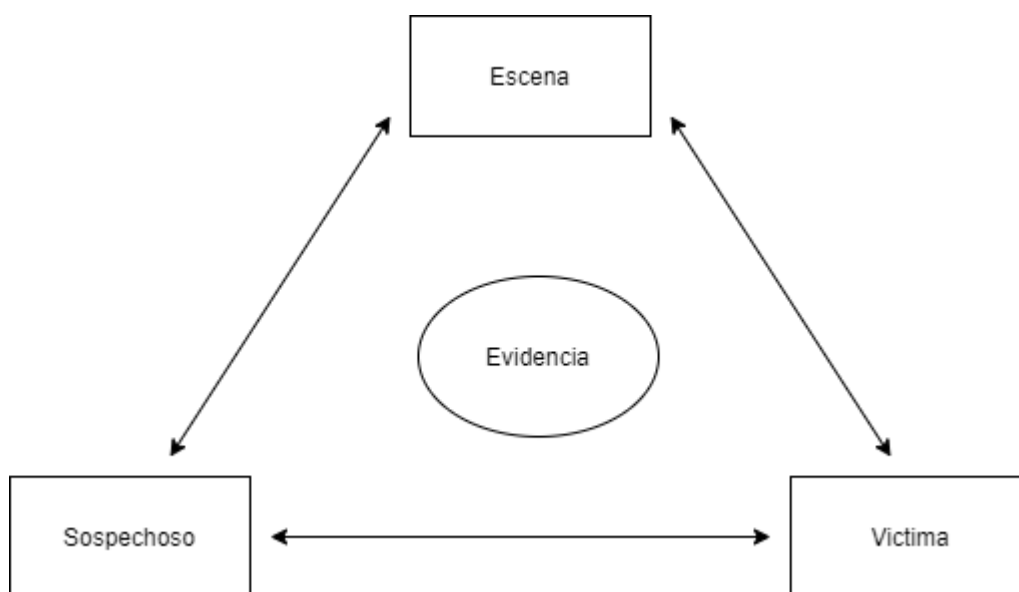


Figure 6: Diagrama del principio de Locard

Los análisis forenses se pueden clasificar de una forma un tanto básica pero efectiva, en función de qué están orientados a analizar, en este caso salen cuatro categorías, que son:

- Análisis de sistemas, dependiendo de si son sistemas GNU/Linux, sistemas macOS y sistemas Windows.
- Análisis de redes.
- Análisis de sistemas empotrados
- Análisis de memorias Volátiles

3.2 Evidencias forenses

En la informática forense, el objetivo siempre es llevar a cabo una recopilación, preservación, análisis y reportes de las posibles evidencias encontradas en una escena de un crimen, para poder posteriormente usarlas como evidencias de procesos legales o como base para una mejoría en un sistema que ha sido vulnerado.

De acuerdo con lo anteriormente explicado, se define una evidencia como cualquier tipo de información extraída de un medio mediante algún tipo de intervención, ya sea humana o automática, de donde deducimos que en el campo de la informática encontramos como algunos ejemplos de evidencia:

- Accesos a ficheros
- Ficheros de monitorización del sistema
- Cookies
- Procesos en ejecución en un sistema
- Memorias físicas

En el campo de las evidencias, las digitales son las que conforman un desafío de mayor calibre por norma general, todo debido a una serie de características de los sistemas informáticos, como son:

- Anonimidad
- Volatilidad
- Redundancia
- Exterminable
- Modificable

Debido a que las evidencias en el caso de los sistemas informáticos se consideran evidencias digitales, se debe considerar que el análisis forense de dichas evidencias, debe ser un análisis forense digital,

3.3 Modelos de informática forense

Dentro de la informática forense no existe un modelo estándar de procedimiento, aun así, existen una serie de modelos a nivel nacional e internacional que pueden servir de guía a la hora de llevar a cabo una investigación forense digital, dentro de los modelos más habituales en el entorno encontramos los explicados en este apartado.

3.3.1 Digital Forensic Research Workshop (DFRW)

El modelo DFRW es un modelo internacional creado por la organización Digital Forensic Research Workshop, en él se hace uso de técnicas que permiten clasificar las diferentes actividades que se llevan a cabo en grupos.

En este modelo se tiende a cubrir etapas de desarrollo en las que se tiene mas en cuenta las etapas que no son de tipo judicial, de esta forma el modelo DFRW queda separado en las siguientes siete fases



Figure 7: Modelo DFRW.

Donde:

1. **Identificación:** Se debe evaluar la situación inicial del crimen, incluidos entorno físico, digital, estado actual del entorno, etc.
2. **Colección:** Se debe recolectar las evidencias del entorno, tanto físico como digital si están disponibles y generar copias.
3. **Preservación:** Se debe garantizar la integridad y confidencialidad de las evidencias y sus copias, además de una cadena de custodia de estas.
4. **Análisis:** En esta fase se debe hacer una consideración de las herramientas y aplicarlas sobre las evidencias.
5. **Examen:** Llegado este punto deberíamos ser capaces de evaluar las evidencias en función de los datos que nos provean.

6. **Informe:** En este paso se debe realizar un informe forense con los datos sacados de las evidencias analizadas para presentar al cliente.
7. **Decisión:** En este último punto se debe dar una decisión final sobre los hechos que han tenido lugar en este caso y en base a las evidencias.

3.3.2 Modelo de análisis forense militar

Una metodología de análisis forense militar posee ciertas diferencias con respecto a los demás modelos de actuación de análisis forense civil, se centra más en aislar los motivos llamaron la atención del atacante a la hora de elegir a la víctima, recuperar el sistema objetivo del ataque y establecer las medidas necesarias para que dicho incidente no se repita.

Cabe destacar que en esta metodología no hace falta la presencia de ninguna entidad judicial, sino que puede contarse con la autoridad militar pertinente para llevar a cabo la tarea, aunque en caso de ser necesario legalmente, el incidente deba siempre ser comunicado a las administraciones correspondientes.

De este modo, el modelo de análisis forense militar queda separado en 5 fases:

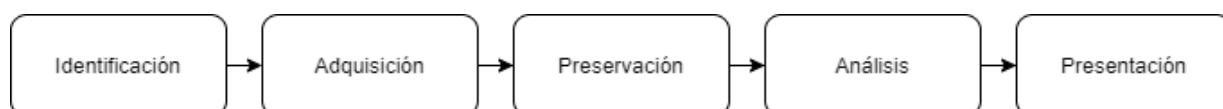


Figure 8: Diagrama del modelo de análisis forense militar.

1. **Identificación:** En esta fase se realiza una evaluación inicial del entorno, se identifica el incidente con su gravedad y se aísla a la víctima.
2. **Adquisición:** Se deben recoger las evidencias haciendo uso de métodos acordes y estableciendo un orden de recolección para las evidencias.
3. **Preservación:** Se hacen copias de seguridad de las evidencias y se establece la cadena de custodia y traslado.
4. **Análisis:** se realiza el análisis técnico de las evidencias haciendo uso de las herramientas y técnicas apropiadas para el tipo de evidencia y se intenta recrear la escena intentando identificar al autor, forma de actuación, etc.

5. **Presentación:** Se realiza una recopilación de todas las evidencias con su información proporcionada y se prepara un informe forense sobre el caso para la autoridad pertinente.

3.3.3 Modelo Casey 2004

El modelo Casey 2004 es una iteración más moderna del anterior modelo Casey 2000, en él se mantiene la misma base, pero debido a que era un sistema evolutivo, se han mejorado y ajustado para tiempos más modernos, el modelo se divide en 8 fases que son las siguientes:

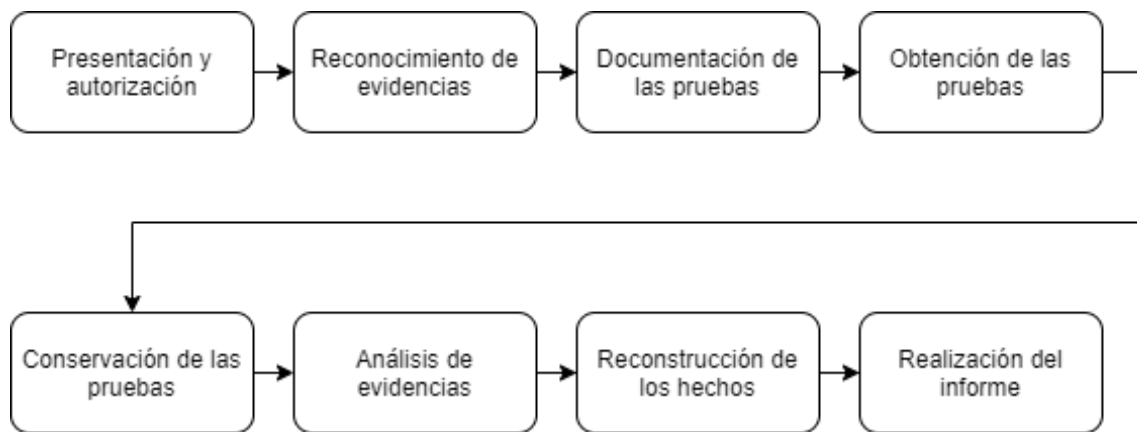


Figure 9: Diagrama Casey 2004.

1. **Preparación y autorización:** Durante esta fase se deben recoger todas las actividades referentes a la recolección de evidencias y presentación ante autoridades legales para su aceptación.
2. **Reconocimiento de evidencias:** Durante el transcurso de esta fase se deben especificar y etiquetar todas las evidencias encontradas.
3. **Documentación de pruebas:** En esta fase se debe realizar una documentación exhaustiva sobre los procesos llevados a cabo para encontrar las pruebas que se usan en este modelo.
4. **Obtención de pruebas:** En esta fase se realizan imágenes de las evidencias digitales encontradas y que vayan a ser analizadas posteriormente, el objetivo es que se pueda utilizar más tarde como prueba en un posible juicio.

5. **Conservación de las pruebas:** Se deben asegurar las pruebas para que no haya ningún problema de contaminación durante el transcurso de esta fase.
6. **Análisis de evidencias:** En esta fase se realiza el análisis de las evidencias digitales encontradas y se deberá formular una hipótesis sobre el caso.
7. **Reconstrucción de los hechos:** Alcanzando esta fase ya se debe poder mediante las pruebas extraídas de las evidencias, realizar una pequeña reconstrucción de los hechos del caso.
8. **Realización del informe:** Siendo la última fase de la metodología lo único que queda es realizar el informe pericial con todo detalle del análisis llevado a cabo.

3.4 Estándares Nacionales

Es necesario para el desarrollo de un análisis forense digital la existencia de una serie de normas, guías o pautas para recolectar, preservar y analizar evidencias digitales de forma correcta, a pesar de la inexistencia a nivel nacional de una regulación estandarizada, existen unos estándares que hacen referencia a la disciplina y que son de ayuda a la hora de llevar a cabo un análisis forense digital.

En este apartado de la memoria del proyecto se hablará de algunos estándares nacionales de especial interés para un análisis forense digital.

3.4.1 Normas UNE 71505:2013 y 71506:2013

Las normas UNE 71505:2013 y 71506:2013 son normas referentes a la gestión de evidencias digitales, su importancia se basa en que hacen referencia a la admisibilidad de las pruebas en juicios.

La norma **UNE 71505:2013** se subdivide en 3 apartados más pequeños representando la estructura que vemos en la siguiente tabla:

Tabla 3.4.1.1: Sub-normas UNE 71505:2013.

UNE 71505-1:2013	En esta sub-norma se hace referencia al vocabulario y principios generales de la norma.
UNE 71505-2:2013	En esta sub-norma se hace referencia a las buenas prácticas en la gestión de evidencias digitales.
UNE 71505-3:2013	En esta sub-norma se tiene en cuenta los formatos y mecanismos técnicos.

Al seguir esta normativa, se garantiza que las evidencias usadas en juicios o disputas legales están correctamente tomadas y gestionadas, de forma que se evita el ser puestas en duda ante cualquier tribunal.

Por otro lado, la norma **UNE 71506:2013** proporciona una serie de buenas prácticas de referencia para la gestión de evidencias electrónicas, además de establecer adicionalmente el análisis de dichas evidencias basándose en los principios de la norma UNE 71505.

3.4.2 RFC 3227

Los RFC son una serie de documentos que dan unas pautas para llevar a cabo un proceso, en el caso del 3227 se indican unas directrices para la recolección de evidencias y como almacenarlas, algunos de los puntos más importantes de este documento son:

- Recolección de evidencias
 - Llevar a cabo la recolección siguiendo el orden de volatilidad de las evidencias, es decir, el periodo en el que son accesibles sin modificación, comenzando por la más volátil.
 - Se debe evitar ciertas acciones a la hora de recolectar las evidencias, como por ejemplo apagar el medio físico en el cual se encuentra la evidencia, no confiar en información de los programas del sistema, etc. Para evitar perder la validez de las evidencias.
 - Es importante que se tenga en cuenta la privacidad de la compañía a la que se le está haciendo el análisis para no invadirla en ningún momento de la recolección de evidencias.
 - La recolección de evidencias debe ser totalmente transparente y seguir unos pasos lo más claros posibles.
- Almacenamiento
 - Se debe establecer una cadena de custodia claramente documentada y en detalle, además, la información se debe almacenar en dispositivos preparados y aptos para el almacenamiento de evidencias forenses.
- Herramientas
 - Las herramientas para recolectar las evidencias deben ser externas al sistema para evitar contaminación de evidencias, deben ser poco intrusivas, etc.

3.4.3 ISO 27042:2015

La ISO 27042:2015 es una norma que trata el análisis y la interpretación de evidencias digitales, para facilitar al investigador, que ante un incidente para el cual se requiere una investigación forense digital, la realización del análisis de la evidencia, desde su identificación, hasta su aceptación como prueba sea correcto.

La norma también ofrece información sobre algunos modelos de análisis que se pueden usar sobre evidencias digitales.

4 Desarrollo de una metodología de análisis forense

4.1 Características y requisitos

Antes de dar una metodología e intentar estandarizar los procesos sobre las bases de datos, se debe establecer una serie de características y requisitos que deben cumplirse para el estudio de este proyecto específico, en este aspecto se encuentra:

1. Las bases de datos objeto de estudio en este proyecto serán relacionales, debido a la vasta aceptación que tienen en el mercado y su amplia extensión e implementación actuales, de forma que las convierte en la mayoría del mercado y por tanto en el público mayoritario a la hora de sufrir incidentes.
2. Es de importancia establecer la ubicación física o remota de la base de datos, puesto que no se actuará de la misma forma en caso de estar ubicada remotamente en un fichero de bases de datos a estar ubicada en un servidor físico con un sistema de gestión de bases de datos.
3. En caso de estar manejada por un SGBD como es habitual, se espera que la base de datos esté ubicada en red, de forma que se debe tener claro si se usa el mismo SGBD a lo largo de toda la red o si se usan distintos SGBDs en la distribución de red.
4. Las bases de datos deben ser dinámicas puesto que los sistemas actuales mayoritariamente están preparados para la modificación de datos, es decir, operaciones de lectura, escritura y borrado.

5. Lo mas habitual es que para realizar la investigación se necesite de un entorno de laboratorio generado específicamente para realizar la investigación, por tanto, se debe emular haciendo uso de máquinas virtuales o máquinas físicas específicamente preparadas.
6. Debe poderse generar un volcado de la base de datos o una extracción del fichero o directorio que la aloje para poder evaluarlo como evidencia.
7. La base de datos debe hacer uso de lenguajes SQL, lo cual no es muy difícil teniendo en cuenta que la gran mayoría hace uso de dicho lenguaje, por esto último se ha decidido estudiarlas en este proyecto.

4.2 Introducción

Una vez estudiados en los anteriores capítulos los modelos y las normas que atañen a esta disciplina dentro de la informática, queda bastante claro que existen una serie de puntos o pautas que son de suma importancia para el desarrollo de una metodología, entre ellos puntos como la conservación del entorno y la identificación de las evidencias correctamente.

Poniendo especial interés en el campo de este proyecto que son las bases de datos y teniendo en cuenta lo anterior, se decide que la mejor forma de establecer una metodología es mediante fases bien diferenciadas, quedando un modelo de 6 fases con la estructura siguiente:

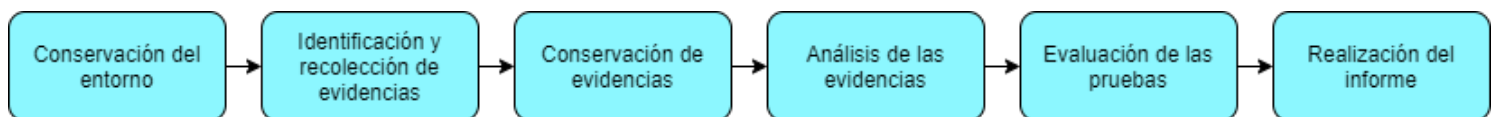


Figure 10: Fases de la metodología.

1. Conservación del entorno.
2. Identificación y recolección de evidencias.
3. Conservación de las evidencias.
4. Análisis de las pruebas.
5. Evaluación de las pruebas.
6. Realización del informe.

Para todas las fases de esta metodología se explicarán las buenas prácticas y puntos más comunes de forma que al hacer una investigación forense digital siguiéndola, haya la menor cantidad de dudas posible y se comentan el mínimo número de errores o imprudencias.

4.3 Fase 1: Conservación del entorno

Esta primera fase se corresponde con el punto 2 de las características y requisitos de este proyecto, el entorno físico en caso de existir debe ser asegurado adecuadamente, en caso de ser remoto se debe investigar todo lo posible sin invadir la privacidad de la empresa que provea el servicio remoto, solicitando en caso necesario el acceso y control completo al entorno remoto

En cualquier caso, que se produzca un incidente, siempre hay un entorno que controlar y conservar, de forma que es importante a la hora de pasar a las fases posteriores, haber asegurado que el entorno no ha sido modificado desde el momento de identificación del incidente y hasta haber realizado el análisis de las evidencias como mínimo, debido a que cualquier modificación del entorno podría variar el estado de una posible prueba del caso.

La recomendación en este sentido es siempre tomar instantáneas del entorno físico, como está todo colocado y montado en el entorno de trabajo, identificando de esta forma el perímetro a analizar y en caso de ser necesario, restringiendo el acceso hasta el final de la investigación.

Se deben proteger además los aspectos físicos del entorno, como por ejemplo posibles huellas dactilares. Además, se debe identificar y etiquetar todo lo que haya podido formar parte del incidente y el momento en el que se etiquetó.

Obviamente a la hora de evaluar un método digital, si la máquina víctima se encuentra encendida se debe preservar su estado encendido y todas las aplicaciones o procesos que puedan estar funcionando en pantalla en ese momento. Las máquinas intervenidas no se deben apagar ni modificar de estado hasta terminada como mínimo la fase de identificación y recolección de evidencias.

4.4 Fase 2: Identificación y recolección de evidencias

Con la creación de esta fase tenemos en cuenta los puntos 1, y 6 de los requisitos explicados en la introducción. Durante esta fase se debe realizar la identificación y recolección de evidencias del entorno previamente asegurado y conservado, dichas evidencias deben ser extraídas de cualquier punto de información que pueda aportar datos relevantes al futuro informe.

Para realizar la identificación se debe primero evaluar el incidente que ha ocasionado todo desde un comienzo, y a partir de ese punto identificar todas las posibles vulnerabilidades o errores que puedan llevar mediante el posterior análisis al hallazgo de una prueba o hecho que ayude a formular una hipótesis de origen del incidente.

Para realizar la correcta recolección de las evidencias se debe realizar copias siempre evitando la modificación de las mismas, lo más apropiado para esto es realizar copias byte a byte de los posibles ficheros, particiones, imágenes que se encuentren y se puedan investigar. Además, estas evidencias deben estar clasificadas dando una serie de datos que ayuden a identificarlas como pueden ser nombre, fecha y hora.

La recolección de las muestras hecha byte a byte, debe además respetar el orden de volatilidad de las evidencias encontradas en el sistema, el cual según la norma RFC 3227, explicado anteriormente en este documento, se especifica como:

1. Registros y contenidos de memoria caché
2. Tabla de enrutamiento y caché ARP, tablas de procesos, estadísticas de procesador, memorias.
3. Información temporal del sistema
4. Discos
5. Logs del sistema
6. Configuración física y topología de red
7. Documentos.

Siguiendo esta cadena de volatilidad de las evidencias existen una serie de acciones que no deben llevarse a cabo antes de la extracción de las evidencias, como son:

- Apagar el medio físico que contenga posibles evidencias.

- Confiar en los programas del sistema, puesto que después del incidente pueden haber sido comprometidos.
- Evitar cambios de fecha y hora en los ficheros del sistema.

4.5 Fase 3: Conservación de las evidencias

Una vez terminada la fase anterior, donde se han identificado y recolectado las evidencias, se debe llevar siempre un control de la manipulación de esas evidencias, puesto que una modificación no deseada de esas evidencias puede invalidarlas completamente ante aspectos legales.

Para evitar problemas, se debe llevar a cabo un control exhaustivo de esas evidencias, para ello se genera una cadena de custodia de las evidencias, cuya finalidad es evitar que se produzcan modificaciones en las evidencias sin tener un control de quien puede haber sido el causante.

La cadena de custodia debe siempre estar claramente documentada y en ella se debe detallar el dónde, quién, cuando y como para todas las posibles cuestiones, como el manejo de la evidencia, la recolección, su almacenamiento y los posibles intercambios o transportes que se realicen con esa evidencia de por medio.

Mientras todas las evidencias estén en un lugar seguro esperando ser analizadas, deben estar embaladas de una forma segura y etiquetadas de forma que se pueda saber que es cada evidencia sin necesidad de desembalarlas, además se deben tener en cuenta posibles fallos eléctricos o magnéticos en posibles componentes físicos que puedan verse afectados por estos.

Por otro lado, debe tenerse muy en cuenta el método de transporte que se usa para las evidencias en caso de tener que transportarlas de un lugar físico al laboratorio donde se realice el análisis, puesto que posibles golpes durante el transporte pueden dejar ciertos volúmenes físicos inservibles, lo cual invalidaría la evidencia.

4.6 Fase 4: Análisis de las pruebas

Al aplicar esta fase tenemos en cuenta especialmente los puntos 3, 4, 5 y 7, puesto que en este punto se debe poder montar un entorno de laboratorio acorde al caso que encontremos para la investigación, esto supone estudiar el tipo de base de datos que se posea del caso específicamente incluyendo su lenguaje, SGBD, modelo de creación, etc.

Durante el transcurso de esta fase se deben evaluar todas las evidencias recolectadas en las anteriores fases para intentar descubrir quien, que o como se causó el incidente que se intenta solucionar. Esta fase finaliza cuando se encuentran pruebas que permiten reconstruir los hechos del incidente o cuando se terminan las evidencias con certeza de que no se van a encontrar más pruebas.

Para poder llevar a cabo un correcto análisis de las evidencias lo primero es tener en cuenta que se deben respetar las leyes actuales a nivel nacional, para poder garantizar que las pruebas que se extraigan de dichas evidencias pueden llegar a ser validas ante un juicio en caso necesario. Los resultados obtenidos del análisis de evidencias deben ser **verificables, reproducibles, independientes** y deben estar correctamente **documentados**.

Para poder comenzar a analizar las evidencias desde un comienzo siempre se debe preparar un sistema de laboratorio que cuente con unas características similares a las condiciones en las que se generó el incidente original, ya sea haciendo uso de máquinas virtuales, máquinas aisladas con características similares, entornos idénticos de sistema operativo, etc. Posteriormente a esta recreación se puede pasar a la visualización de los archivos, sistema de registros, etc. Todo lo que se pudiera extraer como evidencia,

En el caso de este proyecto, se centra en sistemas de bases de datos, por tanto, en la siguiente tabla se ve una serie de herramientas especialmente orientadas hacia análisis forense de bases de datos que pueden ser de ayuda para extracción de evidencias de sistemas de bases de datos.

Tabla 3.4.3.1: Algunas herramientas para realizar análisis sobre bases de datos.

Herramienta	descripción	DBMS
SQL Recon	Escáner de descubrimiento (activo y pasivo) para identificación de servidores Microsoft SQL server en red,	Microsoft SQL server
Recovery for Mysql	Software especialmente diseñado para recuperación de registros y datos almacenados en bases de datos dañadas o corruptas.	Mysql
Acronis Recovery for MS SQL Server	Software de copias de seguridad y recuperación para bases de datos y servidores completos basados en Microsoft SQL server	Microsoft SQL server
Systools SQL recovery	Programa usado para la recuperación de bases de datos completas en lenguaje SQL, con ella se permite la recuperación de datos, tablas, vistas, etc.	Diversos
Stellar Repair for SQLite	Herramienta de reparación de bases de datos, se puede usar para recuperar instancias de la base de datos y posteriormente regenerar la base de datos.	SQLite
Undark	Herramienta que se encarga de parsear toda la base de datos y extraer todos los datos que se encuentren disponibles, eliminados, corruptos o existentes.	SQLite
Forensic toolkit for SQLite	Grupo de aplicaciones para recuperación, investigación y evaluación de bases de datos, permite recuperar registros, reconstruir bases de datos y evaluar imágenes de bases de datos, entre otras cosas.	SQLite

Para casos en los que haya diversos tipos de evidencias, se puede hacer uso de herramientas más comunes, como por ejemplo herramientas de descubrimiento de red, herramientas de recuperación de registros, etc. Todo dependiendo del sistema en el que se vaya a realizar el análisis de evidencias.

Adicionalmente en esta fase y en casos que lo requieran, puede ser necesario realizar una estimación del impacto que ha tenido el incidente en la compañía, especialmente para casos judiciales, donde debería ser necesario poseer datos económicos del coste que ha supuesto el incidente.

4.7 Fase 5: Evaluación de las pruebas

En la quinta fase el objetivo es revisar el análisis de las pruebas y su documentación, de forma que se consiga formular una hipótesis o se consiga asegurar que en base a esas pruebas se puede determinar la causa del incidente, sus autores y cualquier otra información de importancia para el caso.

Para esto siempre se debe actuar con cautela, debido a que, si no se tiene un 100% de seguridad sobre los hechos, todo son conjeturas, siempre hay que estar seguro de las técnicas que se han usado, su veracidad y de que las pruebas no hayan sido manipuladas. En el caso de que las fases anteriores hayan sido respetadas, bien documentadas y llevadas a cabo, se podría llegar a hallar pruebas concluyentes sobre un sospechoso en el caso o sobre el incidente en sí, a pesar de que no siempre se va a obtener una respuesta.

El objetivo de esta fase es intentar reconstruir los hechos del incidente en base a las pruebas encontradas o extraídas de las evidencias, para poder posteriormente adjuntarlas al informe y pueda ser usado en una disputa legal en caso de que sea necesario y se tenga la certeza necesaria sobre los hechos.

4.8 Fase 6: Realización del informe

En esta última fase lo que resta es realizar la redacción del informe forense que acompaña al caso, en él debe incluirse una serie de datos que son:

- Antecedentes del caso: se debe explicar el entorno encontrado al revisar el caso y lo que ha llevado ahí, que puede ser desde los indicios de posibles problemas en las compañías, los reportes de incidentes, etc.
- Metodología seguida: se debe dar un detalle de la metodología a seguir para investigar el caso y el por qué se elige.
- Recolección y gestión de evidencias: se debe indicar las evidencias que se recogen, forma de recolección, identificación de las evidencias, etiquetaje de las evidencias, etc.
- Análisis de las evidencias: se deben detallar, con el nivel técnico necesario dependiendo de si es un informe judicial o no, los procedimientos seguidos para analizar las evidencias encontradas en la recolección y los resultados hallados de esas evidencias.
- Conclusiones del informe: se deben dar unas conclusiones claras y concisa, halladas de las evidencias y sin suposiciones ni sugerencias de solución, puesto que es un informe forense, no de auditoria.

Se debe tener mucho cuidado con el grado de tecnicidad del informe, puesto que, en caso de requerirse un informe para la vía judicial, el nivel técnico del informe debe ser más bajo, puesto que el perfil del usuario que va a dar uso al informe por norma general no es técnico y no poseerá el nivel de conocimiento requerido. En esta serie de casos puede ser recomendable escribir varios informes adaptándose al nivel técnico del lector, aunque no sea un requerimiento.

En cualquier caso, el informe debe ser claro, conciso y directo, debe evaluar las evidencias siguiendo la metodología, se debe documentar correctamente y en caso de tener certeza sobre el origen de los hechos o del sospechoso, apuntarlo directamente.

Este informe es ideal para el sistema que se intenta investigar, puesto que en él se define absolutamente todo, desde lo que ha llevado a la necesidad de realizar la investigación con los antecedentes del caso, hasta las conclusiones del informe halladas a partir de una recolección y análisis de evidencias correctamente guiado, documentado

y apoyado por una metodología también detallada para que haya el menor margen de error.

5

Caso Práctico

A lo largo de este capítulo se expondrá un caso práctico ficticio creado para la aplicación de la metodología desarrollada en el anterior capítulo, estará específicamente centrado en el segmento de las bases de datos.

5.1 Caso práctico

Para el caso práctico de este proyecto se ha creado un caso ficticio, proponiendo la invención de una PYME, puesto que es la mejor orientación para este tipo de proyectos debido a su número y especialmente su baja formación por norma general en el campo de la Informática.

Para este trabajo se desarrolla la empresa “Artencuentro”, una compañía dedicada a la comunicación entre profesionales del mundo del arte y empresas. Como muchas otras PYMEs, se ha creado un sistema web sencillo pero robusto y que funciona en remoto, cuya base de datos se encuentra diseñada en SQLite, el porqué de esta elección se basa en el concepto de que muchas empresas de pequeño alcance hacen uso de CMS para la programación de sus servicios web, debido a su bajo coste, alta capacidad de personalización y rapidez de despliegue.

Para comenzar el caso se da la situación de que la empresa contacta con el investigador debido a unos supuestos errores que existen en el sistema, los cuales están relacionados con la base de datos de la aplicación y que parecen no tener solución para los programadores de la compañía.

Se tiene en cuenta también que la compañía ha puesto en marcha un servicio de una de las copias de seguridad con la configuración en funcionamiento, por tanto, no tienen intención de pasar a la vía judicial, solamente se requiere recabar información sobre el problema, su posible origen y en la medida de lo posible, el autor del problema.

5.2 Preparación del caso práctico

Para preparar el caso práctico se ha necesitado montar el sistema web desde un comienzo, para ello se ha montado una aplicación web haciendo uso de un servidor apache con XAMPP, siendo un modelo cliente-servidor, la programación se ha llevado a cabo en PHP para el lado del servidor y HTML + CSS en la parte del cliente.

El servicio web cuenta con un sistema de Roles, donde existen un administrador, una serie de artistas y unas empresas (véase figura 11 – Anexo 1). A través del servicio web, los artistas pueden subir sus obras con una serie de datos como el título, tipo de obra, fecha de creación, descripción y la obra en sí (Véase figura 12 – Anexo 1). Adicionalmente se habilita una sala de mensajería y proposiciones entre los artistas y las empresas, para que en el caso de que una empresa se vea interesada en una obra/artista, pueda contactar con él en caso de tener algún tipo de oferta (véanse figuras 13 y 14 – Anexo 1).

La base de datos del servicio web se monta en SQLite, puesto que es una base de datos, ligera, sencilla y muy extendida a nivel global, además de que muchos CMS la usan por defecto. El esquema de la base de datos relacional se puede encontrar en la figura 16, bajo el anexo 2. Hace uso de lenguaje SQL y al ser una base de datos SQLite se basa en un sistema de ficheros para su almacenaje, de forma que se crea un fichero “datos.db” dentro del proyecto que corresponde con la base de datos.

Para la exposición del problema se ha intentado llevar a cabo un ataque posible dentro del mundo de la Informática, en la actualidad alto porcentaje de los ataques realizados corresponden con ataques de tipo “ransomware”, el objetivo de estos ataques es llevar a cabo el cifrado de algunos archivos o incluso sistemas completos con la intención de solicitar un rescate por ellos posteriormente. Por tanto, para este proyecto ficticio se ha llevado a cabo un cifrado de la base de datos, simulando una intrusión en el sistema remoto contratado por la compañía y que les habría dejado con la base de datos cifrada.

Para llevar a cabo el cifrado del fichero de la base de datos en SQLite se ha hecho uso de una máquina virtual con sistema operativo Ubuntu que ha realizado una conexión al servicio web, se ha colado haciendo un ataque de tipo “directory traversal” y ha usado la librería OpenSSL para cifrar el fichero “datos.db” con algoritmo RSA (4096).

5.3 Aplicación de la metodología

Para aplicar la metodología en este caso práctico ficticio, se van a encontrar una serie de problemas, que se irán viendo por fases, aunque se detallarán en el próximo apartado, al realizar el informe.

1. En la Fase 1 se debe asegurar y conservar el entorno, el problema que se encontrará es que es un entorno remoto, por tanto, no se tiene acceso físico a los servidores, de forma que no se tiene total control sobre el sistema donde se tiene alojado el servidor web, y por ende solamente se puede llevar a cabo un estudio del proyecto con el servicio web y el sistema de bases de datos.
2. Durante la fase 2 se deben detallar las evidencias extraídas, en este caso será el servicio web y la base de datos, debido a la escasa existencia de evidencias encontradas en el entorno, en este caso será relativamente sencillo detallarlas y etiquetarlas correctamente, además de tener en cuenta la cadena de custodia de las evidencias tomadas.
3. En el transcurso de la fase 3 se debe llevar a cabo la conservación de las evidencias, es decir, de la unidad usada para transportarlas hasta el entorno de laboratorio donde se va a trabajar y su almacenamiento, no debería haber problema puesto que no hay demasiadas evidencias ni son muy difíciles de transportar.
4. Durante la fase 4 se realiza el análisis de las evidencias, en este apartado es donde mas problemas vamos a encontrar, puesto que las herramientas mas habituales de uso para realizar investigaciones forenses de SQLite no están preparadas para lidiar con un fichero de base de datos cifrado, por tanto van a saltar errores de fichero mal formateado o incorrecto, para ello tendremos que pasar a investigar el fichero con un editor de texto y posteriormente con una herramienta que nos indique si de verdad se encuentra cifrado, esta fase se detalla mejor en el informe forense del caso del apartado siguiente.
5. En la fase 5, el investigador se da cuenta de que, al estar el sistema cifrado con una clave bastante potente de cifrado, se deduce que no se puede descifrar de forma útil y se pasa directamente a las conclusiones, se deduce que el fichero ha sido cifrado, posiblemente por algún método externo y que no se puede operar más con él.

6. En esta ultima fase 6, se realiza el informe y se presenta, de forma que el cliente pueda tener claro lo que se ha hecho, como se ha hecho y como se ha llegado a las conclusiones.

5.4 Informe Forense del caso

5.4.1 Antecedentes del caso

El 16 de junio de 2020 se recibe la llamada de una compañía llamada “Artencuentro”, se ha detectado un fallo en su servicio web el cual ha hecho que su servicio haya estado funcionando incorrectamente por un periodo de casi 8 horas con sus correspondientes perdidas. Se quiere saber cuál ha sido la causa del problema y si es posible, su origen.

“Artencuentro” es una compañía que se dedica a la intermediación entre profesionales de artes gráficas y empresas, siendo una pequeña empresa dentro del marco nacional se reducen a un servicio web bastante sencillo al cual nos han dado todo el acceso posible.

El servicio web está montado en PHP con base de datos en SQLite, el problema para hacer el análisis en este caso reside en que el servicio se encuentra totalmente externalizado a la compañía, es decir, no se posee un servidor físico en el cual se encuentra montado el portal web y al cual se pueda acceder para revisar en busca de evidencias.

Desde la compañía se proporciona acceso al servidor remoto con credenciales (que no se expondrán en este documento por privacidad), se proporciona el mensaje de error que dio el indicio del incidente, Un diagrama relacional de la base de datos y una descripción básica del funcionamiento del servicio web, adjuntados en el anexo 1 de este documento.

La compañía “Artencuentro” pretende que se realice una investigación forense sobre el servicio web y sus componentes para descubrir el fallo, de donde pudo haber venido, y con qué objetivo, por tanto, no existe motivación judicial detrás del informe.

La metodología seguida para la realización de este informe ha sido la expuesta en el capítulo 4 de este informe, basada en fases de forma que a la finalización de las 6 fases se pueda intentar esclarecer luz sobre este caso.

5.4.2 Recolección y gestión de evidencias

Para este caso se encuentran una serie de problemas a la hora de realizar la recolección y gestión de evidencias, el sistema se encuentra totalmente externalizado, es decir, en la oficina de la compañía no se encuentra ningún servidor físico al que se tenga acceso y que aloje el servicio web, por otro lado, la externalización del servicio no nos permite acceso a los registros del sistema, tampoco a las tablas de enrutamiento ni discos y logs del sistema.

Lo único de lo que se dispone para trabajar el sistema de ficheros donde se encuentra alojado el proyecto de la aplicación web y la base de datos, así que se opta por recolectarlos realizando una copia byte a byte mediante SSH y se guardan en una memoria externa identificada y embalada correctamente para su transporte al laboratorio de pruebas.

El sistema remoto no se encontraba apagado a la hora de realizar las pruebas así que el sistema sigue en el mismo estado exactamente que cuando el incidente ocurrió.

En la cadena de custodia de las evidencias se encuentra:

Tabla 5.4.2.1: cadena de custodia de evidencias.

Evidencia	Recolección	Fecha y hora	Encargado
Sistema de ficheros de la aplicación	Mediante SSH	16/06/2020 17:35	Eduardo Rodríguez Hernández
Base de datos de la aplicación SQLite	Mediante SSH	16/06/2020 17:40	Eduardo Rodríguez Hernández

5.4.3 Análisis de las evidencias encontradas

A pesar de haber encontrado escasas evidencias debido al sistema de la compañía, lo cual dificulta el hecho de encontrar posibles problemas, se dispone a investigar las evidencias encontradas.

Se comienza por montar el entorno de laboratorio para funcionamiento, se monta con un servidor apache basado en XAMPP en una máquina limpia con sistema operativo Windows 10 en su ultimo nivel de actualizaciones, exactamente igual que el del sistema remoto que la compañía Artencuentro tiene contratado.

Una vez el sistema está funcionando, se encuentra con que el servicio web tira de una base de datos montada en SQLite, cuyo fichero de almacenamiento se denomina “datos.db” también extraído, por tanto, se comienza con las pruebas.

En primera instancia se encuentra que el error persiste en el servicio web, es decir, se encuentra el mismo error que en la figura 15 del anexo 2 proporcionada por la compañía y que indica que existe algún tipo de fallo en la base de datos, por tanto, se dispone a revisar dicho fichero.

La primera herramienta usada para investigar la base de datos es “Stellar Repair for SQLite” la cual debería permitir llevar a cabo una visualización y reparación de la base de datos, pero el resultado por parte de la aplicación es negativo, indicando que el tipo de fichero es incompatible como se ve en la figura 17 del anexo 3.

Se intenta por otro lado con la herramienta “SysTools SQLite Database Recovery”, la cual, en caso de corrupción del fichero de la base de datos, permitiría recuperarlo o regenerar una nueva base de datos con los registros recuperados de la original además de permitir ver posibles problemas con la base de datos. Nuevamente esta aplicación también nos devuelve resultado negativo, el fichero no tiene un formato valido como se ve en la figura 18 del anexo 3.

Ante las 2 negativas de aplicaciones de renombre, solo queda investigar directamente el fichero en busca de un posible fallo en el formato del fichero de base de datos de SQLite, al utilizar una aplicación de revisión de ficheros, en este caso “binwalk” en su versión para Windows se nos da como resultado que el fichero se encuentra cifrado (véase figura 19 en el anexo 3).

Al encontrarse que el fichero de la base de datos está cifrado y que su entropía es bastante elevada, damos por sentado que el fichero ha sido cifrado por algún método posiblemente externo a la compañía (véase figura 20 del anexo 3).

5.4.4 Conclusiones del informe

Con la información y las evidencias recabadas a lo largo de este caso, se deduce que el incidente se generó por un cifrado no deseado en el fichero. Posiblemente se trate de un ataque de ransomware en el cual se intenta cifrar la base de datos del

servicio como sistema básico y se solicita un rescate por él, pero a falta de pruebas ante accesos al servicio, Logs de sistema, registro de Windows, etc. No se puede deducir con total certeza que se encuentre este caso.

Se devuelven las pruebas a la compañía “Artencuentro” con la cadena de custodia y todo lo pertinente para que, si se desea, se solicite a la compañía que oferta el hosting del servicio control total de la máquina y llevar a cabo un informe forense adicional que esclarezca mas hechos sobre este caso.

6 Conclusiones y líneas futuras

En este último capítulo de la memoria se hará un breve repaso al proyecto y se darán posibles líneas futuras de trabajo que puedan derivar de este proyecto, también se darán algunos problemas encontrados a lo largo del desarrollo del proyecto y unas conclusiones al proyecto.

6.1 Repaso y conclusiones

Como se ha visto a lo largo del desarrollo de esta memoria, no existe una estandarización al respecto de procedimientos forenses en el mundo de la Informática, solamente existen una serie de guías y buenas pautas proporcionadas por entidades estatales y privadas sobre cómo realizar un análisis forense, lo cual deja a cada investigador trabajando de una forma personal y probablemente distinta de uno a otro.

La falta de estandarización no implica el hecho de que las investigaciones no se puedan realizar con total confianza, pero sí indica que se deben tener una serie de puntos en cuenta siempre y que puede que a veces se pasen por alto.

Especialmente teniendo en cuenta el caso de las bases de datos y su importancia en el mercado actual, puesto que todos los sistemas informáticos se soportan sobre alguna base de datos de algún tipo, pero a la hora de hacer investigaciones forenses sobre posibles corrupciones de bases de datos, apenas existe información al respecto.

El acercamiento de este trabajo era desarrollar una metodología que, aunque de carácter genérico, pueda servir específicamente para ayudar a terceros a realizar investigaciones forenses sobre bases de datos de una forma un tanto más guiada y sencilla.

Para ello se propone la metodología del capítulo 4 de esta memoria, la cual en el desarrollo del caso práctico de este trabajo ha ayudado a realizar un informe relativamente acorde a los estándares legales y de formato, aun siendo un caso complicado debido al nivel de evidencia con el que se contaba y el problema que eso genera.

En base a las consideraciones del caso práctico, cabe destacar que, por razones de extensión y tiempo, ha sido imposible realizar un examen mucho más exhaustivo y en diferentes entornos o con más casos prácticos, los cuales permitirían llevar a cabo un examen de efectividad de la metodología.

6.2 Líneas futuras de trabajo

Como líneas futuras a este trabajo se podría proponer:

- Un examen más exhaustivo de la metodología generada, incluyendo variaciones a la misma en caso de encontrar alguna posible mejora y diversos casos de estudio los cuales ayudarían a probar la efectividad de dicha metodología.
- Se podría plantear generar guías más precisas para la presentación a nivel legal de un informe forense digital, puesto que es algo realmente complicado y que puede influir de sobremanera en un caso judicial real.
- Se podría probar más exhaustivamente la efectividad de la metodología contra la efectividad de otras metodologías existentes a nivel mundial.
- Se podría proponer el desarrollo de algún sistema a nivel más técnico que ayude o aporte al mundo del análisis forense en bases de datos, debido a la gran escasez que hay de herramientas y guías sobre operaciones posibles.

7

Anexos

7.1 Anexo 1: capturas del servicio web.



The screenshot displays the 'Artencuentro' web application. At the top, the title 'Artencuentro' is underlined in blue. Below it is a black navigation bar with the text 'Ver nuestros trabajos' on the left, 'Elma Ndamás' and 'Logout' in the center, and a search bar on the right. The main content area features a table with user information. The table has two columns: 'nombre' and a column for actions. The users listed are Elma Ndamás, Artis Taenun Saco, Elme Jordi Bujante, Pint Orde Brocha, Construcciones Eternas, and Mark Etingcan Arias. Each user row has 'Editar' and 'Eliminar' links. At the bottom of the table is a 'Crear Usuario' link.

nombre	
Elma Ndamás	Editar Eliminar
Artis Taenun Saco	Editar Eliminar
Elme Jordi Bujante	Editar Eliminar
Pint Orde Brocha	Editar Eliminar
Construcciones Eternas	Editar Eliminar
Mark Etingcan Arias	Editar Eliminar
Crear Usuario	

Figure 11: Sistema de roles.


Artencuentro					
titulo	tipo	fecha	descripcion	imagen	
obratfm	impresionista	03/06/2020	una obra para una prueba del tfm obviamente		Artis Taenun Saco

Figure 12: Obra de un artista.

Artencuentro					
nuestros trabajos			Ver mis propuestas	Mark Etingcan Arias	Logout
Autor	hora	descripcion	presupuesto		
Pint Orde Brocha	2018-03-06 11:57:28	Escuela de Ingenieria Informatica: desarrollar letras y logo de 4 metros de largo a situar en el exterior de edificio	2500	Ver Mensajes	

Figure 13: ejemplo de propuesta.

hora	mensaje	Remitente
	veo, que le interesa la propuesta, estaremos encantados de darle mas informacion	Mark Etingcan Arias
1970-01-01 00:00:17	me interesa	Pint Orde Brocha
<input type="text" value="Escriba su mensaje"/> <input type="button" value="Enviar"/>		

Figure 14: Ejemplo de mensajería.

7.2 Anexo 2: mensajes de error, diagrama de la base de datos del caso, descripción del servicio.

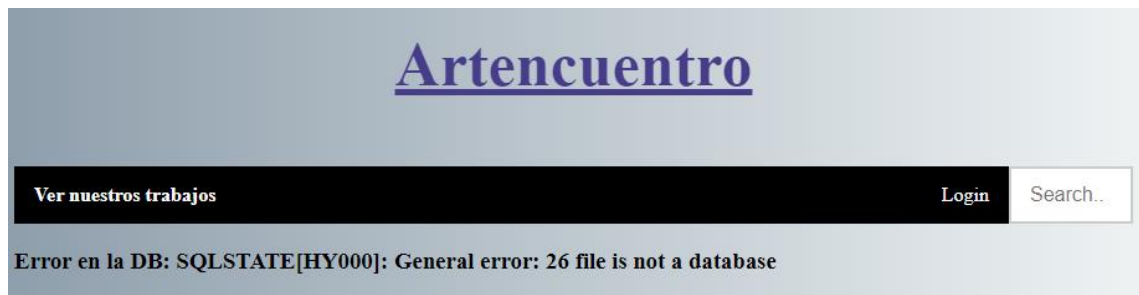


Figure 15: Error en la base de datos.

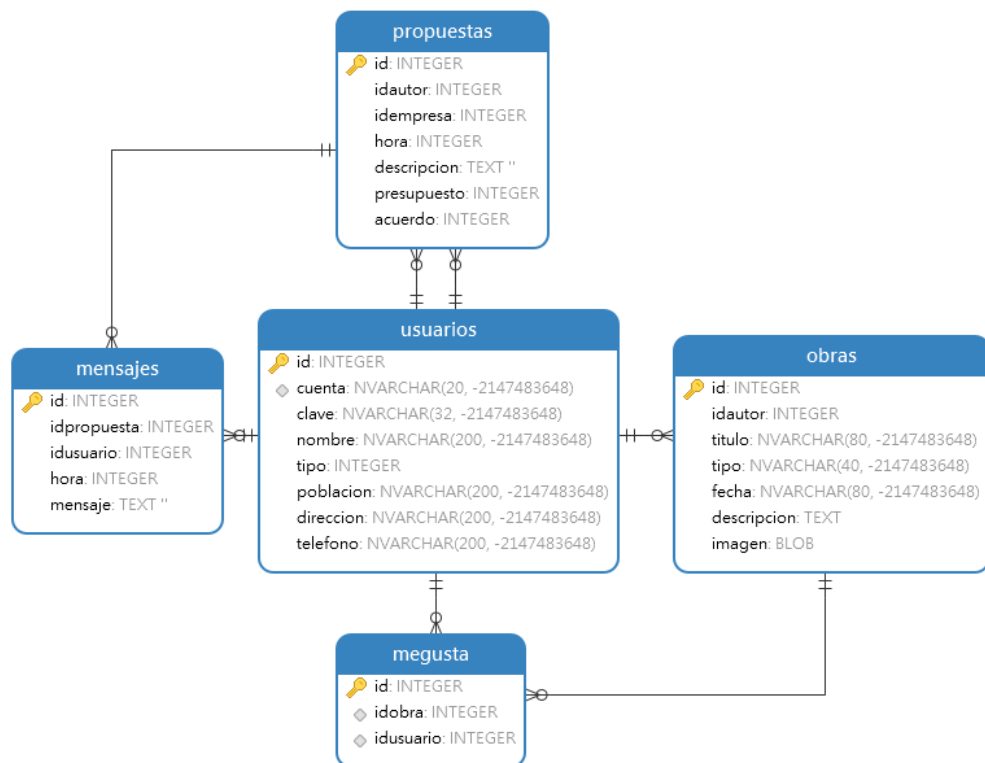


Figure 16: Diagrama de la base de datos.

A modo de explicación se recibe un pequeño resumen de la aplicación que contiene los siguientes puntos:

- “Artencuentro” es un servicio web que permite a los artistas exponer su trabajo en internet. Existe un sistema de roles que diferencia entre administrador, artista y empresa.
- Los artistas tienen total control sobre sus obras, pueden subirlas, eliminarlas y modificarlas si desean.
- Las empresas pueden ver las obras de los artistas, recibir información sobre ellos y ponerse en contacto con los artistas mediante un sistema de mensajería entre usuarios.
- El administrador posee todo el poder sobre el servicio web y los usuarios.
- La base de datos está montada en SQLite y el servicio en PHP + CSS + HTML.

7.3 Anexo 3: Capturas de pantalla

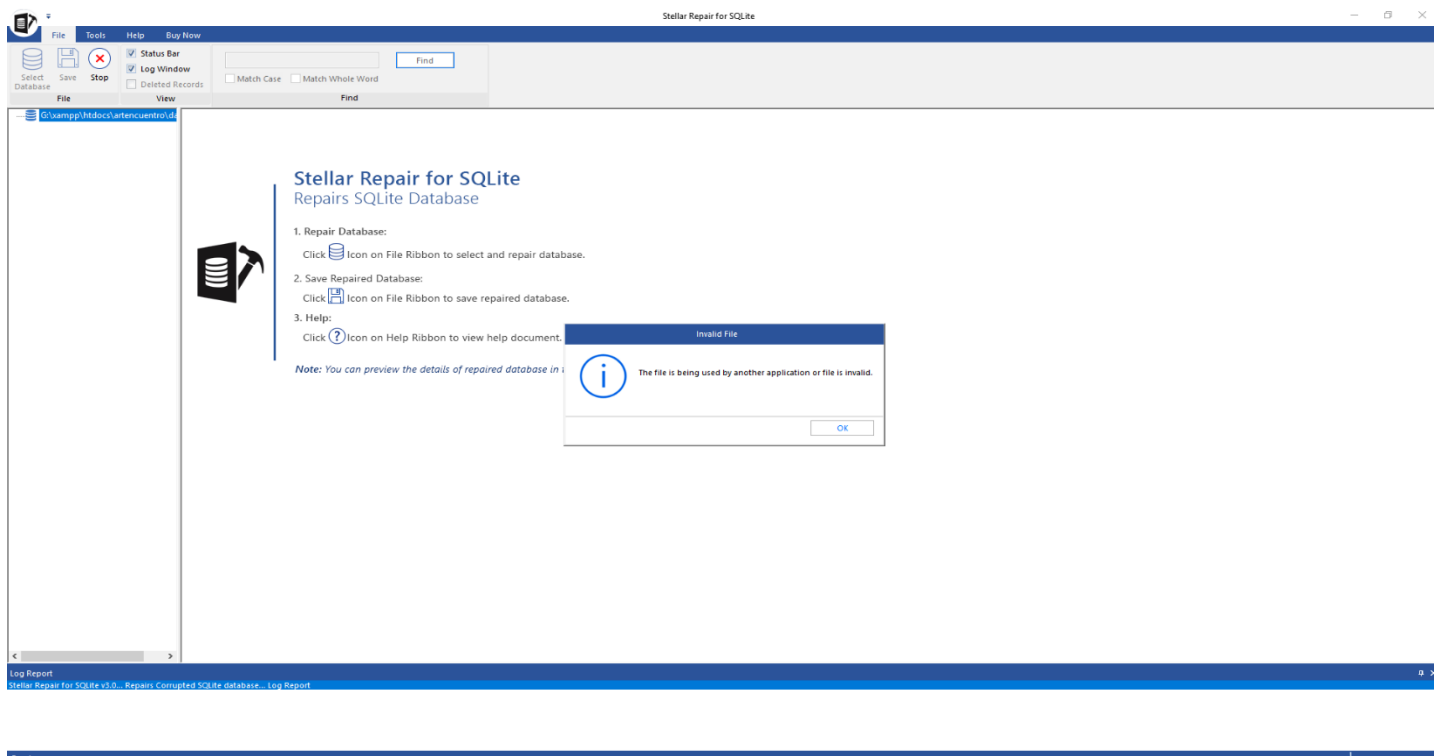


Figure 17: Stellar Repair for SQLite.

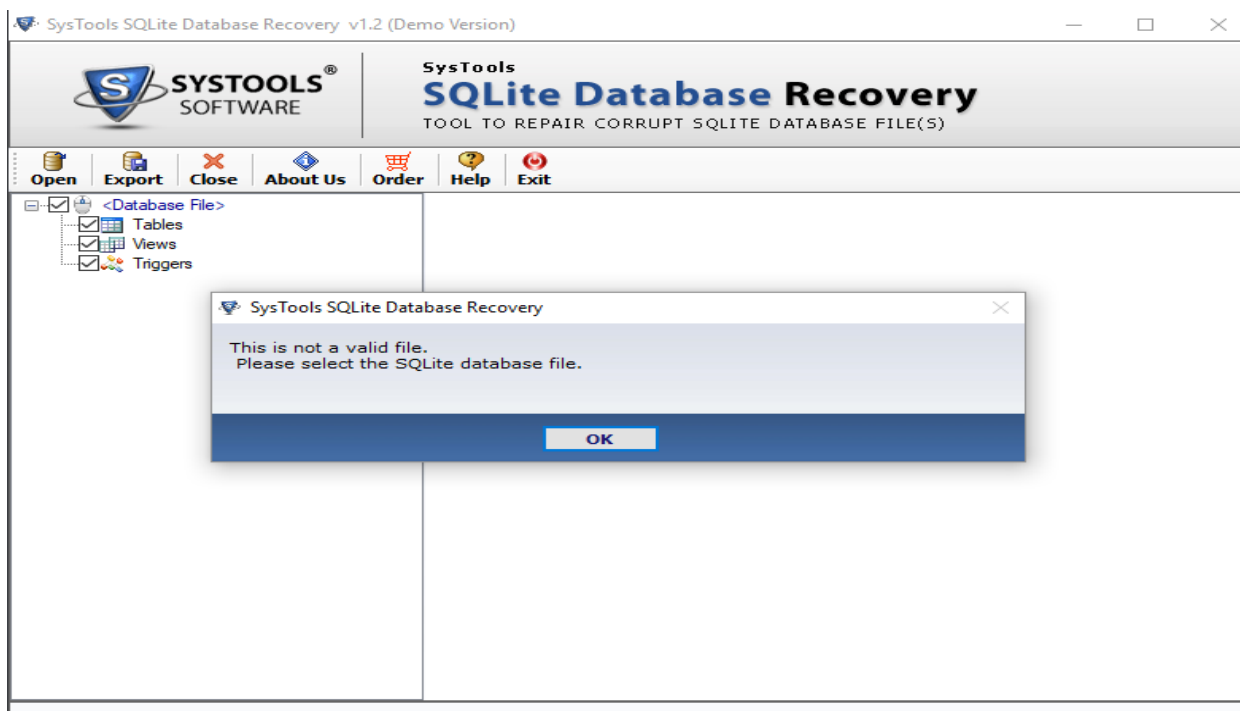


Figure 18: SysTools SQLite Database Recovery.

```

edu@DESKTOP-RPPG7HR:/mnt/g/xampp/htdocs/artencuentro$ binwalk datos.db

```

DECIMAL	HEXADECIMAL	DESCRIPTION
0	0x0	OpenSSL encryption, salted, salt: 0xC3FB1BAE28792717

Figure 19: Binwalk sobre base de datos.

```

edu@DESKTOP-RPPG7HR:/mnt/g/xampp/htdocs/artencuentro$ binwalk -E datos.db

```

DECIMAL	HEXADECIMAL	ENTROPY
0	0x0	Rising entropy edge (0.972968)
1891328	0x1CDC00	Falling entropy edge (0.617188)

Figure 20: entropía del fichero de bases de datos.

8

Bibliografía

- [1] INCIBE, «RFC 3227 - Directrices para la recopilación de evidencias y su almacenamiento,» 2014.
- [2] Ninjas de la web, «Metodología de Análisis Forense,» [En línea]. Available: <https://ninjasdelaweb.com/metodologia-de-analisis-forense/>. [Último acceso: 25 may 2020].
- [3] INVACI, «Diez estándares y guías para el análisis, investigación y peritaje en informática y telemática forense,» [En línea]. Available: <https://javiermarques.es/diez-estandares-informatica-forense>. [Último acceso: 25 may 2020].
- [4] Informatico Forense, «Modelos de análisis informático forense,» [En línea]. Available: <https://www.informatico-forense.es/modelos-de-analisis-informatico-forense/>. [Último acceso: 27 April 2020].
- [5] Portaltic, «Qué es la informática forense y cómo se usa para resolver casos policiales o judiciales,» [En línea]. Available: <https://www.europapress.es/portaltic/sector/noticia-informatica-forense-usa-resolver-casos-policiales-judiciales-20181124112932.html>. [Último acceso: 27 April 2020].
- [6] INCIBE, «Guía de toma de evidencias en entornos windows,» November 2014. [En línea]. Available: https://moodle.upm.es/titulaciones/oficiales/pluginfile.php/7960465/mod_resource/content/1/incibe_toma_evidencias_analisis_forense.pdf.
- [7] INESEM, «<https://revistadigital.inesem.es/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>,» 16 April 2019. [En línea]. Available: <https://revistadigital.inesem.es/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>. [Último acceso: 20 April 2020].
- [8] Ciberseguridad.com, «Análisis forense,» 2019. [En línea]. Available: <https://ciberseguridad.com/servicios/analisis-forense/>. [Último acceso: April 2020].
- [9] E. G. Castro, *Guía de Actuación de un Ingeniero Forense en el ambito de fallos de materiales.*, Sevilla: Universidad de sevilla., 2018.

- [1] J. R. Alamillo, «Estándares nacionales e internacionales que puede seguir un perito informático para realizar el análisis forense de una evidencia y para la elaboración de un peritaje informático,» november 2016. [En línea]. Available: <https://peritoinformaticocolegiado.es/blog/estandares-nacionales-e-internacionales-que-puede-seguir-un-perito-informatico-para-realizar-el-analisis-forense-de-una-evidencia-y-para-la-elaboracion-de-un-peritaje-informatico/>. [Último acceso: 2020].
- [1] unknown, «Seguridad en las bases de datos».
- [1] SQLite (company), «How To Corrupt An SQLite Database File,» [En línea]. Available: <https://www.sqlite.org/howtocorrupt.html>. [Último acceso: April 2020].
- [1] J. Wagner, "https://www.youtube.com/watch?v=l0ecuo8nGks," in *DePy*, 2016.
- [1] S. N. F. F. Sven Schmitt, «A Standardized Corpus For SQLite Database Forensics,» de *DFRWS EU*, Florence, 2018.
- [1] Forensic Focus, «Forensic Analysis of Damaged SQLite Databases,» [En línea]. Available: <https://www.forensicfocus.com/articles/forensic-analysis-of-damaged-sqlite-databases/>. [Último acceso: April 2020].
- [1] INCIBE, «¿Quieres trabajar en informática forense?. Estos son los principales dominios de conocimiento.,» INCIBE, 2013. [En línea]. Available: <https://www.incibe-cert.es/blog/dominios-de-conocimiento-informatica-forense>. [Último acceso: April 2020].
- [1] G. Messina, "What Is Database Forensics?," INFOSEC, [Online]. Available: <https://resources.infosecinstitute.com/category/computerforensics/introduction/areas-of-study/application-forensics/overview-types-of-database-forensics/>. [Accessed May 2020].
- [1] Wikipedia, «Database Forensics,» December 2019. [En línea]. Available: https://en.wikipedia.org/wiki/Database_forensics. [Último acceso: March 2020].
- [1] «Sistema de gestión de bases de datos,» 6 June 2020. [En línea]. Available: https://es.wikipedia.org/wiki/Sistema_de_gestión_de_bases_de_datos. [Último acceso: June 2020].