

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA



TESIS DE GRADO
MODELO DE DETECCIÓN DE IRRUPCIONES EN
INFORMÁTICA FORENSE

PARA OPTAR AL TÍTULO DE LICENCIATURA EN INFORMÁTICA
MENCIÓN: INGENIERÍA DE SISTEMAS INFORMÁTICOS

AUTOR : Univ. Javier Quispe Tambo
TUTOR : Lic. Eufren Llanque Quispe
REVISOR: Lic. Javier Reyes Pacheco

La Paz – Bolivia

2009



DEDICATORIA

Dedico la presente Tesis, a mi madre Susana Tambo Vda., de Quispe por su apoyo constante e incondicional. A mi padre Emilio Quispe (+), por su apoyo en los inicios de mi carrera universitaria. A mis hermanos (as) por el incentivo y colaboración.



AGRADECIMIENTOS

A Dios por la infinita bendición, por guiar mis pasos por el sendero del bien de la cultura y la educación.

A todas las autoridades, docentes, compañeros (as) de la carrera, Facultad de Ciencias Puras y Naturales.

Un agradecimiento especial a mi revisor y tutor de la presente Tesis, por las enseñanzas, comprensión, apoyo, aliento que me brindaron en los momentos difíciles que me tocó vivir.

RESUMEN

La informática Forense es un campo nuevo en la investigación informática, de tal forma la presente tesis de grado enfoca su estudio a la detección de irrupciones dentro la informática forense.

Se describe una breve introducción acerca del tema para contextualizar, posteriormente se presenta el problema de investigación, también se muestran el objetivo general como guía del estudio además de sus objetivos específicos que apoyan al logro de la detección de irrupciones en informática forense; posteriormente se formula una hipótesis.

Contiene la teoría acerca de las irrupciones, tipos de irrupciones, definiciones sobre la informática forense, los pasos normales que se desarrollan en la investigación forense, también describe los tipos de ataques que podría existir en la informática, se menciona la conceptualización de criptografía. Cabe hacer notar que se resalta la teoría sobre los modelos ocultos de Markov.

Presenta la investigación de la detección de irrupciones en informática forense. Lo primero que se realiza es desarrollar las fases para la detección de irrupciones en informática forense los cuales son: fase de Validación de usuarios, identificación, preservación de evidencias, análisis de evidencias, informes. Y posteriormente, se aplica los modelos ocultos de Markov para la elaboración del prototipo, para finalmente, realizar las pruebas y resultados que verifican el análisis de sensibilidad.

Finalmente, describe las conclusiones y recomendaciones a las cuales se arribaron en la investigación.

ÍNDICE GENERAL

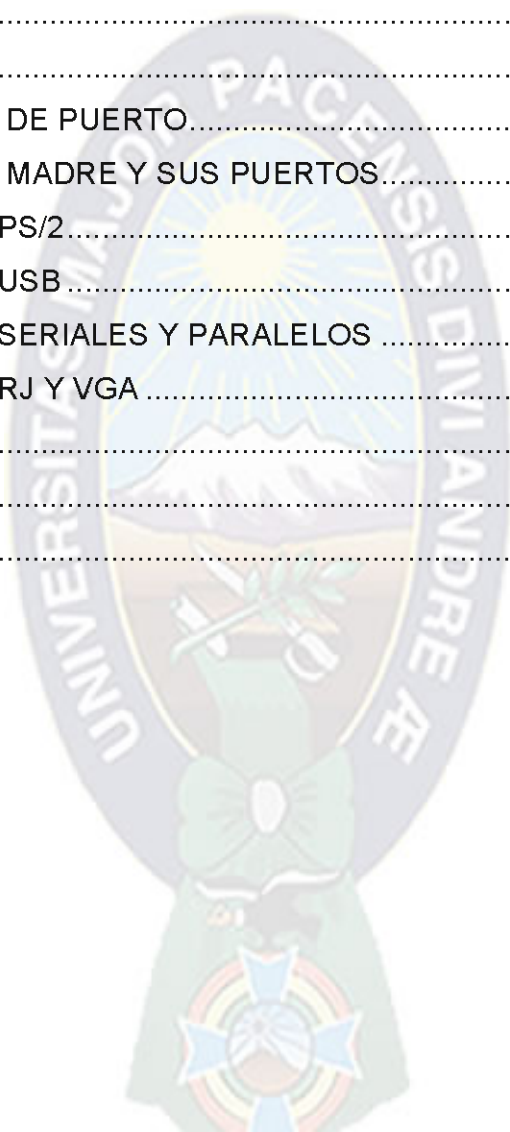
| | |
|---|----|
| CAPÍTULO 1 | 1 |
| ASPECTOS GENERALES | 1 |
| 1.1. INTRODUCCIÓN..... | 1 |
| 1.2. ANTECEDENTES | 2 |
| 1.3. PLANTEAMIENTO DEL PROBLEMA..... | 3 |
| 1.3.1. Descripción..... | 3 |
| 1.3.2. Formulación de la Pregunta | 4 |
| 1.4. OBJETIVOS | 4 |
| 1.4.1. Objetivo General..... | 4 |
| 1.4.2. Objetivos Específicos | 4 |
| 1.5. FORMULACIÓN DE HIPÓTESIS | 5 |
| 1.6. JUSTIFICACIONES..... | 5 |
| 1.6.1. Justificación Teórica | 5 |
| 1.6.2. Justificación Metodológica..... | 5 |
| 1.6.3. Justificación Económica | 6 |
| 1.6.4. Justificación Técnica | 6 |
| 1.6.5. Justificación Social | 6 |
| 1.7. METODOLOGÍA Y HERRAMIENTAS | 6 |
| 1.7.1. Método científico | 7 |
| 1.8. ALCANCES Y APORTES..... | 9 |
| 1.8.1. Alcances..... | 9 |
| 1.8.2. Aportes | 9 |
| CAPÍTULO 2 | 10 |
| MARCO TEÓRICO..... | 10 |
| 2.1. LA CIENCIA FORENSE | 10 |
| 2.2. DEFINICIÓN DE IRRUPCIONES | 13 |
| 2.2.1. Definición y significado de Interrupciones | 14 |
| 2.3. ATAQUE DE LOS VIRUS..... | 15 |
| 2.3.1. Identificación del ataque..... | 15 |

| | | |
|----------|---|----|
| 2.3.2. | Descripción del Virus | 15 |
| 2.3.3. | Evaluación de Antecedentes | 17 |
| 2.3.4. | Identificación y Evaluación del Vector Infeccioso | 18 |
| 2.3.5. | El Virus penetra en la red | 18 |
| 2.3.6. | Esfuerzos | 20 |
| 2.3.7. | Contención del ataque..... | 22 |
| 2.4. | TIPOS DE DETECCIÓN..... | 24 |
| 2.4.1. | Fuentes de información Forense | 26 |
| 2.4.1.1. | NIDS..... | 26 |
| 2.4.1.2. | HIDS..... | 28 |
| 2.5. | CÓMPUTO FORENSE | 32 |
| 2.5.1. | Dispositivos a Analizar | 33 |
| 2.5.2. | Pasos del cómputo forense | 34 |
| 2.5.2.1. | Identificación | 34 |
| 2.5.2.2. | Preservación | 34 |
| 2.5.2.3. | Análisis | 34 |
| 2.5.2.4. | Presentación | 36 |
| 2.5.3. | Finalidad de la Informática Forense | 36 |
| 2.5.4. | Objetivos de la informática Forense | 37 |
| 2.6. | TÉCNICAS ANTI FORENSE | 38 |
| 2.7. | SEGURIDAD INFORMÁTICA..... | 40 |
| 2.7.1. | Objetivos | 43 |
| 2.7.2. | Análisis de Riesgos | 44 |
| 2.7.3. | Elementos de un Análisis | 45 |
| 2.7.4. | Análisis de Impacto al Negocio | 46 |
| 2.7.5. | Puesta en Marcha de una Política de Seguridad | 47 |
| 2.7.6. | Las Amenazas..... | 48 |
| 2.7.7. | Tipos de amenazas | 49 |
| 2.7.8. | La amenaza Informática del futuro | 50 |
| 2.7.8.1. | En el futuro | 51 |
| 2.7.9. | Técnicas para asegurar el sistema..... | 52 |

| | | |
|-----------|--|----|
| 2.7.9.1. | Consideraciones de <i>Software</i> | 52 |
| 2.7.9.2. | Consideraciones de una red..... | 53 |
| 2.7.10. | Afirmaciones Erróneas de Seguridad..... | 53 |
| 2.7.10.1. | Mi sistema no es importante para un <i>cracker</i> | 53 |
| 2.7.10.2. | Estoy protegido pues no abro archivos que no conozco..... | 54 |
| 2.7.10.3. | Como tengo antivirus estoy protegido..... | 54 |
| 2.7.10.4. | Como dispongo de un firewall no me contagio..... | 54 |
| 2.7.10.5. | Servidor <i>Web</i> actualizado a la fecha..... | 54 |
| 2.7.11. | Organismos Oficiales de seguridad Informática..... | 55 |
| 2.7.12. | Costos elevados en la Seguridad Total..... | 55 |
| 2.8. | MODELO OCULTO DE MÁRKOV..... | 55 |
| 2.8.1. | Historia..... | 57 |
| 2.8.2. | Arquitectura de un modelo oculto de Markov..... | 58 |
| 2.8.3. | Probabilidad de una secuencia observada..... | 58 |
| 2.8.4. | Definición formal de un Modelo Oculto de Markov..... | 59 |
| 2.8.5. | Aplicaciones Modelos Ocultos de Markov..... | 60 |
| 2.9. | CRIPTOGRAFÍA..... | 61 |
| 2.9.1. | Conceptos..... | 61 |
| 2.9.2. | Historia de la criptografía..... | 63 |
| 2.9.3. | Ramas derivadas..... | 66 |
| 2.9.4. | Algoritmos..... | 66 |
| 2.9.5. | Protocolos..... | 67 |
| 2.9.6. | Aplicaciones..... | 67 |
| 2.10. | PUERTOS..... | 68 |
| 2.10.1. | Introducción..... | 68 |
| 2.10.2. | Puerto lógico..... | 68 |
| 2.10.3. | Puerto Físico..... | 69 |
| 2.11. | LEYES..... | 73 |
| 2.11.1. | Ley 1768 del Código Penal, 2 artículos de Delitos Informático..... | 73 |
| 2.11.2. | Reglamento de soporte lógico, D.S. 24582..... | 74 |

| | |
|--|-----|
| CAPÍTULO 3 | 84 |
| ELABORACIÓN DEL MODELO DE DETECCIÓN DE IRRUPCIONES | 84 |
| 3.1. DEFINICIÓN DE FASES DEL MODELO..... | 84 |
| 3.1.1. Validación de usuarios | 85 |
| 3.1.2. Identificación de Evidencias | 88 |
| 3.1.3. Preservar las Evidencias | 91 |
| 3.1.4. Análisis de las Evidencias | 96 |
| 3.1.5. Presentación é Informes..... | 100 |
| 3.2. DIAGRAMA GENERAL | 101 |
| 3.2.1. Cadena de comportamiento Prioritario..... | 101 |
| 3.2.2. Comportamiento en la Red..... | 102 |
| 3.3. MODELADO MEDIANTE MARKOV | 102 |
| 3.3.1. Modelos Ocultos de Markov | 103 |
| 3.3.2. Generación de Observaciones | 104 |
| 3.3.3. Representación del diagrama de Estado..... | 105 |
| 3.3.4. Proceso de los Modelos Ocultos de Markov | 105 |
| 3.3.5. Cómputo de $P(O \lambda)$ | 106 |
| 3.3.6. El Algoritmo de Avance | 107 |
| 3.3.7. Ilustración del Algoritmo de avance..... | 107 |
| 3.3.8. Algoritmo de Retroceso | 108 |
| 3.3.9. Ilustración del procedimiento de Retroceso..... | 109 |
| 3.3.10. Secuencias Óptimas de Estado..... | 109 |
| 3.3.11. Algoritmo de <i>Viterbi</i> | 110 |
| 3.3.12. Algoritmo de Reestimación de <i>Baum-Welch</i> | 111 |
| 3.4. PROTOTIPO | 113 |
| 3.5. PRUEBAS Y EXPERIMENTACIÓN..... | 119 |
| 3.5.1. Definición De Variables De Prueba | 119 |
| 3.5.2. Datos para la Experimentación..... | 119 |
| 3.5.3. Verificación de la Hipótesis | 123 |
| 3.5.4. Resultados | 123 |

| | |
|---|-----|
| CAPÍTULO 4 | 124 |
| CONCLUSIONES Y RECOMENDACIONES | 124 |
| 4.1. CONCLUSIONES | 124 |
| 4.2. RECOMENDACIONES..... | 125 |
| REFERENCIAS BIBLIOGRÁFICAS | 126 |
| BIBLIOGRAFÍA | 126 |
| WEBGRAFÍA..... | 128 |
| ANEXO 1: NÚMEROS DE PUERTO..... | 130 |
| ANEXOS 2: TARJETA MADRE Y SUS PUERTOS..... | 135 |
| ANEXO 3: PUERTOS PS/2..... | 136 |
| ANEXO 4: PUERTOS USB | 137 |
| ANEXO 5: PUERTOS SERIALES Y PARALELOS | 138 |
| ANEXO 6: PUERTOS RJ Y VGA | 139 |
| ANEXO 7: RCA | 140 |
| ANEXO 8: GPS | 141 |
| ANEXO 9: YANAPTI | 142 |



ÍNDICE DE TABLAS

| | |
|---|-----|
| Tabla N° 1: Normas reglamentarias en Bolivia..... | 3 |
| Tabla N° 2: Identificación de Variables..... | 119 |
| Tabla N° 3: Pruebas para el experimento 1..... | 120 |
| Tabla N° 4: Pruebas para el experimento 2..... | 120 |
| Tabla N° 5: Pruebas para el experimento 3..... | 120 |
| Tabla N° 6: Pruebas para el experimento 4..... | 121 |
| Tabla N° 7: Pruebas para el experimento 5..... | 121 |
| Tabla N° 8: Resultados de Prueba..... | 121 |



ÍNDICE DE FIGURAS

| | |
|--|-----|
| Figura N°1: Seguridad y detección de Intrusos | 27 |
| Figura N°2: Arquitectura | 58 |
| Figura N°3: Máquina Enigma Utilizada por los alemanes durante la II guerra Mundial..... | 65 |
| Figura N°4: Fases de la detección de Irrupciones..... | 85 |
| Figura N°5: Validación de Usuario | 86 |
| Figura N°6: Planteamiento del problema a resolver..... | 87 |
| Figura N°7: Identificación de Evidencias | 90 |
| Figura N°8: Preservar Evidencias | 92 |
| Figura N°9: Duplicado a nivel bit | 93 |
| Figura N°10: Empaquetado de dispositivos..... | 93 |
| Figura N°11: Análisis de Evidencias..... | 98 |
| Figura N°12: Presentación de Informes..... | 100 |
| Figura N°13: Cadena de comportamiento | 101 |
| Figura N°14: Comportamiento de la Red..... | 102 |
| Figura N°15: Transiciones | 104 |
| Figura N°16: Transiciones | 105 |
| Figura N°17: Transiciones | 105 |
| Figura N°18: Ilustración del algoritmo de Avance..... | 107 |
| Figura N°19: Ilustración del procedimiento de Retroceso..... | 109 |
| Figura N°20: Procedimiento de Reestimación..... | 112 |
| Figura N°21: Interfaz Inicial | 114 |
| Figura N°22: Selección de documento para la detección de Irrupción | 114 |
| Figura N°23: Archivo Seleccionado | 115 |
| Figura N°24: Verificación de Vulnerabilidad | 116 |
| Figura N°25: Detección de Irrupción en el puerto..... | 116 |
| Figura N°26: Clonación de Archivo | 117 |
| Figura N°27: Analizando Evidencias | 117 |
| Figura N°28: Informe de Evidencia..... | 118 |

| | |
|--|-----|
| Figura N°29: Puertos detectados | 118 |
| Figura N°30: Cuadro estadístico de Aciertos y desaciertos..... | 122 |
| Figura N°31: Porcentaje de Aciertos | 122 |
| Figura N°32: Resultados en Porcentaje | 122 |



CAPÍTULO 1 ASPECTOS GENERALES

1.1.INTRODUCCIÓN

El constante reporte de vulnerabilidades en sistemas de información, el aprovechamiento de fallas bien sea humanas, procedimentales o tecnológicas sobre infraestructuras de computación en el mundo, ofrecen un escenario perfecto para que se cultiven tendencias relacionadas con intrusos informáticos. Estos intrusos poseen diferentes motivaciones, alcances y estrategias que desconciertan a analistas, consultores y cuerpos de especiales de investigaciones, pues sus modalidades de ataque y penetración de sistemas varían de un caso a otro.

A pesar del escenario anterior, la criminalística ofrece un espacio de análisis y estudio hacia una reflexión profunda sobre los hechos y las evidencias que se identifican en el lugar donde se llevaron a cabo las acciones catalogadas como criminales. En este momento, es preciso establecer un nuevo conjunto de herramientas, estrategias y acciones para descubrir en los medios informáticos, la evidencia digital que sustente y verifique las afirmaciones que sobre los hechos delictivos se han materializado en el caso bajo estudio.

La informática forense hace entonces su aparición como una disciplina auxiliar de la justicia moderna, para enfrentar los desafíos y técnicas de los intrusos informáticos, así como garante de la verdad alrededor de la evidencia digital que se pudiese aportar en un proceso.

En consecuencia, este breve documento busca ofrecer un panorama general de esta especialidad técnico-legal, para ilustrar sobre los fundamentos generales y bases de actuación de aquellos que se han dedicado a procurar el esclarecimiento de los hechos en medios informáticos, unos nuevos científicos que a través de la formalidad de los procesos y la precisión de la técnica buscan decirle a los intrusos informáticos que están preparados para confrontarlos y procesarlos.

1.2. ANTECEDENTES

Un virus informático es un malware que tiene por objeto alterar el normal funcionamiento de la computadora, sin el permiso o el conocimiento del usuario. Los virus, habitualmente, reemplazan archivos ejecutables por otros infectados con el código de este. El primer virus que atacó a una máquina IBM Serie 360 (y reconocido como tal), fue llamado *Creeper*, creado en 1972. Existen diversos tipos de virus, varían según su función o la manera en que éste se ejecuta en la computadora alterando la actividad de la misma, entre los más comunes están: los gusanos, las bombas lógicas y los troyanos.

Un troyano tiene la función de robar información o alterar el sistema del *Hardware* o en un caso extremo permite que un usuario externo pueda controlar el equipo. Entre los troyanos están los *Spywares* o archivos espías son unas diminutas aplicaciones cuyo objetivo es el envío de datos del sistema donde están instalados, acceden al sistema sin que el usuario sea directamente consciente de ello y se ejecutan en segundo plano.

Con los primeros *Spywares*, nacen las primeras irrupciones a sistemas, obteniendo de esta manera datos del usuario como: páginas visitadas, tiempo de navegación, transferir datos como direcciones de correo electrónico y en el peor de los casos la obtención de documentación privada del usuario.

En lo que se refiere a estudios realizados Detección de Irrupciones en Informática Forense, en el medio son escasos y los pocos que se conocen son más bien planteamientos de tipo teórico, por ser una de las áreas de reciente exploración en el mundo informático.

Una de las investigaciones más recientes sobre el tema la realizó el Ing. Guido Rosales Uriona de la Empresa Yanapti (*Security your e-live*) enfocándose en dos áreas de esta especialidad las cuales son: "La Evidencia Digital" tomando puntos

como puntos importantes los delitos informáticos y los comunes, y “La Informática Forense” desarrollando temas como: El marco legal, Los servicios profesionales, Formación, Recursos de infraestructura, *Hardware y Software*, Casos reales en Bolivia.

En Bolivia existe un Marco Legal sobre Delitos Informáticos los cuales se detallan a continuación:

| Fecha de Promulgación | Descripción |
|-----------------------|---|
| 11/03/97 | Ley 1768 del Código Penal, 2 artículos de Delitos Informáticos |
| 25/04/97 | Reglamento de soporte lógico, D.S. 24582 |
| 01/04/98 | Ley 1836 del Tribunal Constitucional ART. Admite Demandas y Recursos por Fax. |
| 02/08/03 | Ley 2492 Código Tributario, reconoce medios de prueba informáticos, ART. 79 Incorpora los medios tecnológicos, aprueba las notificaciones electrónicas. |
| 01/07/04 | Resolución de Directorio BCB N° 086/2004 aprueba reglamento de Firma Digital. |

Tabla N° 1: Normas reglamentarias en Bolivia.
Fuente: [HON-2009]

1.3. PLANTEAMIENTO DEL PROBLEMA

1.3.1. Descripción

El área de Informática Forense, es un campo muy amplio se pensó en estudiar uno de sus componentes como es forense en redes, siendo este también un escenario muy complejo, pues es necesario comprender la manera como los protocolos, configuraciones e infraestructuras de comunicaciones se conjugan para dar como resultados un momento específico en el tiempo y un comportamiento particular.

El problema que existe en la realidad es no poder obtener información fidedigna sobre las irrupciones en el sistema operativo, que se constituyan en una prueba para sancionar este delito informático.

1.3.2. Formulación de la Pregunta

¿Cómo realizar un modelo de detección de irrupciones por un agente ajeno al sistema, en un tiempo determinado?

1.4. OBJETIVOS

1.4.1. Objetivo General

Diseñar un modelo que permita detectar irrupciones en un determinado lapso de tiempo, para obtener evidencia digital sobre actos fraudulentos.

1.4.2. Objetivos Específicos

- Asegurar mediante este modelo, la obtención de información confiable, que se constituya en evidencia de peso para llegar a la verdad sobre los hechos fraudulentos.
- Desarrollar e investigar sobre los procedimientos para obtener las evidencias necesarias para llevar a cabo la justicia.
- Proporcionar información sobre manipulación informática y accesos no autorizados que ayude a resolver casos con gran dificultad.
- Avanzar sobre la investigación de este tema en nuestro país, que se constituya un aporte para próximos estudios.

1.5.FORMULACIÓN DE HIPÓTESIS

“El Modelo de Detección de Irrupciones, realiza el análisis sobre los puertos donde se producen actos fraudulentos como el acceso no autorizado y la manipulación de Información”.

1.6.JUSTIFICACIONES

1.6.1. Justificación Teórica

El presente trabajo busca realizar un aporte en el Área de Redes Informáticas, en el campo de Informática Forense, pretendiendo obtener un conjunto de técnicas, procedimientos y metodologías aplicables a la detección de Irrupciones de Sistemas.

Se procura ser base de otros trabajos relacionados con el área, abrir una brecha para el desarrollo de nuevos conocimientos y herramientas que exploten los procesos existentes sobre la Informática Forense.

1.6.2. Justificación Metodológica

La tecnología tiene un acelerado crecimiento en lo que se refiere al *Hardware* y *Software*, esto conlleva grandes facilidades de procesar cantidades inmensas de información y lograr cálculos que no se lograrían de manera manual. Esta ventaja origina la utilización de la computadora para representar el comportamiento de sistemas complejos como ser las metodologías para detectar irrupciones en el sistema.

1.6.3. Justificación Económica

La tecnología permite realizar experimentos utilizando la computadora como herramienta y ahorrando de esta manera recursos y tiempo tomando en cuenta el costo horas hombre. Realizar el modelo sin la ayuda de un ordenador involucraría la construcción de dispositivos que representen una ayuda, para realizar este objetivo.

1.6.4. Justificación Técnica

Las exigencias de un mundo basado en el uso de la tecnología de la información, hace posible manejar información en gran cantidad y por su importancia esta capta la atención de personas inescrupulosas quienes buscan sacarle beneficios malintencionados como realizar estafas electrónicas, este es el motivo por el cual se busca realizar mecanismos de Detección de Irrupciones al Sistema.

1.6.5. Justificación Social

La utilidad en este sentido es significativamente importante, puesto que con este modelo se podrá realizar el estudio el acceso no autorizado y la manipulación de información y las consecuencias de estos, por lo que es importante que las autoridades competentes tomen decisiones y políticas para el futuro.

1.7. METODOLOGÍA Y HERRAMIENTAS

El desarrollo de la Tesis de Grado sigue el Método Científico, el cual permite recorrer el camino de la investigación de manera razonable y productiva de forma que se alcancen los objetivos planteados.

1.7.1. Método científico

Cada ciencia, y aun cada investigación concreta, generan su propio método de investigación. Como método de forma general se entiende el proceso mediante el cual una teoría científica es validada o bien descartada. La forma clásica del método de la ciencia ha sido la inducción (formalizada por Francis Bacon en la ciencia moderna), pero que ha sido fuertemente cuestionada como el método de la ciencia, especialmente por Karl Popper, quien sostiene que el método de la ciencia es el hipotético-deductivo.

En todo caso, cualquier método científico requiere estos criterios:

La reproducibilidad, es decir, la capacidad de repetir un determinado experimento en cualquier lugar y por cualquier persona. Esto se basa, esencialmente, en la comunicación y publicidad de los resultados obtenidos. En la actualidad éstos se publican generalmente en revistas científicas y revisadas por pares.

La falsabilidad, es decir, la capacidad de una teoría de ser sometida a potenciales pruebas que la contradigan. Bajo este criterio se delimita el ámbito de lo que es ciencia de cualquier otro conocimiento que no lo sea: es el denominado criterio de demarcación de Karl Popper. La corroboración experimental de una teoría científicamente "probada" —aun la más fundamental de ellas— se mantiene siempre abierta a escrutinio.

En las ciencias empíricas no es posible la verificación; no existe el "conocimiento perfecto", es decir, "probado". En las ciencias formales las deducciones lógicas o demostraciones matemáticas, prueban solamente dentro del marco del sistema definido por unos axiomas y unas reglas de inferencia; el sistema lógico perfecto, que sería consistente, decidible y completo, no es posible, según el teorema de Gödel.

Existe una serie de pasos inherentes al proceso científico, pasos que suelen ser respetados en la construcción y desarrollo de nuevas teorías.

Éstos son:

- Observación: consiste en el registro de fenómenos que forman parte de una muestra.
- Descripción: trata de una detallada descripción del fenómeno.
- Inducción: la extracción del principio general implícito en los resultados observados.
- Hipótesis: planteamiento de las hipótesis que expliquen dichos resultados y su relación causa-efecto.
- Experimentación: comprobación de las hipótesis por medio de la experimentación controlada.
- Demostración o refutación de las hipótesis
- Comparación universal: constante contrastación de hipótesis con la realidad.

Las herramientas a utilizar se detallan a continuación:

- ❖ Técnicas de Ingeniería de *Software* como ser:
 - a) Recopilación de información de personas expertas en el área de la Informática Forense y Forense en Redes, mediante encuestas y entrevistas directas.

- b) Recopilación de información y bibliografía acerca de las Técnicas de Detección, su construcción y sus aplicaciones.

❖ La sistematización de la información recopilada:

- a) Creación de ficheros bibliográficos.
- b) Selección de la información a ser utilizada.

❖ Implementación de un prototipo, que demuestre de manera clara la eficiencia y/o eficacia de la simulación.

1.8. ALCANCES Y APORTES

1.8.1. Alcances

El presente trabajo pretende establecer un modelo basado en procedimientos normados en el país el cual realice la investigación sobre estos delitos para aportar información que conduzca a la reducción de la delincuencia informática.

Este trabajo, no pretende alcanzar una detección a accesos no autorizados en todos los sistemas operativos, sino que será de uso exclusivo para el Sistema Operativo *Windows XP*.

1.8.2. Aportes

Con la realización del presente trabajo se pretende realizar una investigación sobre uno de los campos de la Informática Forense como es la detención de accesos no autorizados que sirva de base para una investigación más completa sobre los campos que abarca la Informática Forense.

CAPÍTULO 2 MARCO TEÓRICO

2.1. LA CIENCIA FORENSE

La ciencia forense es metódica y se basa en acciones premeditadas para reunir pruebas y analizarlas. La tecnología, en caso de análisis forense en sistemas informáticos, son aplicaciones que hacen un papel importante en reunir la información y pruebas necesarias. La escena del crimen es el ordenador y la red a la cual éste está conectado.

El objetivo de un análisis forense informático es realizar un proceso de búsqueda detallada para reconstruir a través de todos los medios, el log¹ de acontecimientos que tuvieron lugar desde el momento cuando el sistema estuvo en su estado íntegro hasta el momento de detección de un compromiso.

Esa tarea debe ser llevada a cabo con máxima cautela, asegurándose que se conserva intacta, a la mayor medida posible, la información contenida en el disco de un sistema comprometido, de forma similar que los investigadores policiales intentan mantener la escena del crimen intacta, hasta que se recogen todas las pruebas posibles.

El trabajo de un investigador forense es necesario para ofrecer un punto de partida fundamental para que los investigadores policiales, ofreciéndoles pistas sólidas, así como pruebas para su utilización posterior.

Cada uno de los incidentes es único, por lo tanto, la involucración de un investigador forense externo es diferente en cada caso. Algunas veces el trabajo puede estar limitado a colaborar con las agencias del gobierno como Departamento de Delitos Telemáticos de Guardia Civil y/o Brigada Investigación Tecnológica,

¹ Log: registro oficial de eventos durante un período de tiempo en particular.

proporcionándoles el equipo íntegro para que sea analizado en sus instalaciones y por sus expertos.

Otras veces realizan una recolección de información del sistema informático: Analizar ficheros log, estudiar el sistema de ficheros² (FS) del equipo comprometido y reconstruir la secuencia de eventos para tener una imagen clara y global del incidente.

El análisis termina cuando el forense tiene conocimiento de cómo se produjo el compromiso, bajo qué circunstancias, la identidad de posible/s atacante/s, su procedencia y origen, fechas de compromiso, objetivos del/los atacante/s así como, cuando ha sido reconstruida completamente la línea temporal de los eventos.

Cuando un investigador forense empieza el análisis de la situación nunca sabe con lo que va a enfrentarse. Al principio puede ser que no encuentre a simple vista ninguna huella ni prueba de que el equipo ha sido comprometido, especialmente si hay un "rootkit"³ instalado en la máquina. Puede encontrar procesos extraños ejecutándose con puertos abiertos. También es frecuente que vea una partición ocupada de su capacidad, pero cuando la verifica a través del sistema muestra otro porcentaje de ocupación. Puede encontrar una saturación de tráfico de red desde un *host*⁴ específico. Es posible encontrar aplicaciones que están consumiendo un porcentaje elevado de del CPU⁵ pero no haya ningún indicio de un programa con ese nombre en el sistema de ficheros.

Los pasos para empezar la investigación de un incidente son diferentes en cada caso. El investigador debe tomar decisiones basándose en su experiencia y el

² Sistema de Ficheros: Conjunto algoritmos y estructuras auxiliares que permitir de manera sencilla y transparente acceder a datos en dispositivos de almacenamiento.

³ Rootkit: Son conjuntos de programas que permiten al mal hechor tomar el control del sistema con todos los privilegios.

⁴ Host: Son los computadores conectados a la red, que proveen y/o utilizan servicios a/de red.

⁵ CPU: Unidad Central de Proceso

"sexto sentido" para llegar al fondo del asunto. No es necesario seguir pasos determinados, el orden no es importante a veces.

Puede que algunos pasos básicos sean más de lo que hace falta y también puede ser que estos sean insuficientes para solucionar el problema. Los pasos básicos pueden concluir en localizar todas las huellas y eventos que se produjeron.

Y en supuestos los pasos básicos no han desvelado la situación, se debe recurrir a llevar a cabo un análisis profundo o de compilación de las aplicaciones encontradas durante la búsqueda. Estas aplicaciones pueden ser escritas totalmente desde cero y protegidas, pero en la mayoría de los casos son aplicaciones utilizadas de forma común, que circulan por la red, estén o no estén protegidas. Cuando hablamos de protección de ficheros podemos hablar sobre técnicas de confusión, ofuscación y compresión.

Utilizando el criterio orden de los procedimientos se establecen por los analistas forenses, se debe considerar como recursos y el orden necesario en cada caso puede variar. Una vez aprendidas técnicas generales, se podrá combinarlos con la experiencia y crear sus propios trucos en un futuro.

La persona que ha descubierto el incidente debe asegurarse que hay máxima información intacta posible para que el investigador forense pueda realizar su trabajo con éxito, ya que la información encontrada dentro del sistema registra la historia real de lo que ha sucedido.

Hay solo única cosa que es común para cada investigación forense, y no es suficiente repetirla siempre. Se debe tener a mano un cuaderno y un bolígrafo para apuntar inmediatamente todos los pasos que efectúa durante el proceso de investigación. También se cuenta con los pasos para preservar y reunir las evidencias; deben ser efectuadas con lentitud, precaución, metódica y pensándolo

dos veces antes de hacer cualquier cosa ya que cualquier error puede llevar consigo consecuencias como pérdida de pruebas.

En el estudio forense tener el cuaderno a mano puede ser necesario para refrescarle la memoria varios meses después de la investigación cuando llegue la hora de testificar en una sala de juicio (si el caso llega a estos extremos) de forma que el informático forense debe tener las evidencias de forma similar al mismo con excepción de tener la información en dispositivos de almacenamiento. Referente a las técnicas de análisis forense descritas a continuación se asume que se utiliza un sistema operativo sobre un dispositivo de almacenamiento.

2.2. DEFINICIÓN DE IRRUPCIONES

También conocidas como IRQ⁶ recursos que utiliza un dispositivo cuando necesita detener el proceso que está realizando la CPU para informar, por su parte está haciendo algo. Si dos dispositivos utilizan la misma interrupción, se produce un conflicto, el ordenador no sabe qué elemento intenta avisarle y suelen aparecer problemas de funcionamiento.

Señal que capta la atención de la CPU y que usualmente se genera cuando se requiere una entrada/salida. Por ejemplo, cuando se presiona una tecla o se desplaza el *mouse*, se generan interrupciones de *Hardware*⁷. Las interrupciones de *Software*⁸ son generadas por un programa que requiere entrada o salida de disco.

Un temporizador interno puede interrumpir continuamente el computador varias veces por segundo, para mantener actualizada la hora o con el propósito de trabajar en tiempo compartido.

⁶ IRQ: Interrupt ReQuest - solicitud de interrupción

⁷ *Hardware*: la parte física del ordenador

⁸ *Software*: la parte lógica del ordenador

Cuando ocurre una interrupción, el control se transfiere al sistema operativo, el cual determina la acción a emprender. Todas las interrupciones tienen prioridades; a mayor prioridad, más rápidamente será atendida la interrupción.

2.2.1. Definición y significado de Interrupciones

- IRQ (*Interrupt Request*). Requerimiento de interrupción. Interrupción de *Hardware* en un computador personal. Ocho líneas (0-7 en 8086/88) y 16 líneas (0-15 en 286 y superiores) aceptan interrupciones de dispositivos.
- NMI (*Non-Maskable Interrupt*) Interrupción *Hardware*: Interrupción prioritaria sobre IRQ e irreversible.
- SAI (Sistema de Alimentación Ininterrumpida) son aparatos que entran en funcionamiento cuando se produce una interrupción en la fuente principal de energía, lo que permite operar durante un tiempo limitado.
- *Streaming* Consiste en una tecnología utilizada para permitir la visualización y la audición de un archivo mientras se está descargando, a través de la construcción de un buffer por parte.
- PIC (*Programmable Interrupt Controller*) Controlador de Interrupciones Programable, se encarga de la comunicación entre los periféricos y el procesador.

2.3. ATAQUE DE LOS VIRUS

2.3.1. Identificación del ataque

El correo electrónico es el método más corriente de contagio de los virus, pero no es la única forma en que éstos pueden penetrar en el entorno de la red IP⁹. Por ejemplo, los usuarios finales pueden traer disquetes infectados de casa, o pueden efectuar descargas FTP¹⁰ o HTTP¹¹ desde sitios infectados. Recientemente, los creadores de virus e intrusos expertos se han puesto de acuerdo para desarrollar códigos virales que penetran las redes aprovechando conocidas fallas de seguridad en varias aplicaciones.

Una vez que un virus penetra en su red, se desplaza de computador en computador de muchas maneras. Algunos virus buscan en la red sistemas configurados para permitir la utilización compartida de archivos y tratan de acceder e infectar los archivos.

Otros virus se envían a sí mismos por el correo electrónico con destino a nodos en la red. Algunos hacen ambas cosas, mientras que otros se propagan por medios inesperados, incluyendo el uso de sistemas de mensajería instantánea o de aplicaciones entre iguales. Un solo computador infectado en su red puede rápidamente infectar muchos otros sistemas en la red IP.

2.3.2. Descripción del Virus

Si el *Software* antivirus puede detectar una infección o un intento de infección, usualmente podrá tratar efectivamente la situación y, por lo tanto, usted no tendrá

⁹ IP: *Internet* Protocol, Una dirección IP es un número que identifica de manera lógica y jerárquica a una interfaz de un dispositivo (habitualmente una computadora) dentro de una red que utilice el protocolo IP.

¹⁰ FTP: File Transfer Protocol usado en *Internet*

¹¹ HTTP: Hypertext Transfer Protocol", en español "Protocolo de Transferencia de Hipertexto"

un incidente de virus. Estos incidentes se producen cuando un virus puede escapar de su programa antivirus y/o la exploración de la detección de intrusos.

Cuando esto ocurre, el virus, típicamente, hará notar su presencia, bien como resultado directo de su intento por propagarse o bien como un efecto secundario. Entre los indicadores usuales de una infección por virus, se encuentran los siguientes:

Sonidos o imágenes inesperados en la pantalla. Especialmente si estos ocurren en múltiples sistemas, pueden ser la señal de la existencia de virus. Aunque estos indicadores no son destructivos, esto no significa que el virus mismo no lo sea.

Los indicadores de archivo son los más comunes, pero generalmente los más difíciles de detectar. Incluyen la aparición de múltiples archivos desconocidos en las estaciones de trabajo del usuario o en los servidores de archivos; la desaparición de múltiples archivos por razones desconocidas; la pérdida de datos dentro de los archivos de datos; o el reemplazo del contenido de los archivos. Si el virus es capaz de infectar archivos, aquellos que contienen códigos ejecutables pueden repentinamente cambiar de tamaño, cuando el virus se inserta a sí mismo en el código y se ejecuta cuando un usuario o una aplicación intentan hacer funcionar el código en el archivo original.

Los indicadores del sistema son usualmente fáciles de detectar puesto que a menudo interfieren con la capacidad para usar el sistema. Entre los ejemplos se incluyen la imposibilidad de dividir los archivos o la destrucción de sistemas completos de archivos. Esta clase de daño sucede rara vez porque interfiere con la capacidad del virus para propagarse. Cuando ocurre, es a menudo el efecto secundario de una pobre programación por parte del creador del virus pero puede ser también el efecto de la así llamada bomba lógica, una porción de código malicioso colocada por el creador para que se ejecute en una fecha específica o se

base en algún otro desencadenante. Sus usuarios le informarán siempre acerca de esta clase de indicador.

Los indicadores de red son provocados usualmente por los efectos secundarios que causan los intentos del virus por propagarse e incluyen tormentas en la red e interrupciones no programadas del correo electrónico. Esta clase de indicador resulta generalmente obvio para muchos usuarios al mismo tiempo, pero puede detectarse también a través del uso de herramientas administrativas de la red con capacidad para dar la alarma.

Los indicadores personalizados son aquellos que usted instala en su propio entorno específicamente para detectar nuevos virus no descubiertos por el *Software* antivirus. Por ejemplo, usted puede querer instalar un grupo de listas de cuentas ficticias para correo electrónico *Microsoft Exchange*, que incluya solamente cuentas de usuarios ficticios, en tal forma que pueda detectar gusanos de correo electrónico que utilizan *Microsoft Outlook* para propagarse.

Mientras a muchos indicadores de virus se les puede seguir fácilmente la pista hasta una acción o un desencadenante específicos, en otros casos el indicador puede presentarse de manera impredecible, al azar. Estos indicadores son los más difíciles de rastrear y de determinar si un virus, en efecto, los está causando.

2.3.3. Evaluación de Antecedentes

El primer paso es erradicar los indicadores no virales. Programas burlones, mensajes publicitarios, errores de aplicación, equivocaciones comunes del usuario, fallas de los sistemas y fallas de *Hardware* de la red que se encuentran entre los eventos que pueden provocar indicadores confusos.

Luego se tiene varias fuentes que pueden ayudarle a identificar los problemas conocidos de tipo viral y no viral que podrían causar los mismos indicadores:

2.3.4. Identificación y Evaluación del Vector Infeccioso

Una vez que se determina que un virus es la causa de un indicador, el siguiente paso es identificar la naturaleza del ataque. Sería ideal que usted tuviera el tiempo suficiente para determinar completamente qué sistemas se hallan afectados, pero algunos virus pueden propagarse más rápido de lo que usted tardaría en evaluar su impacto. Las infecciones que no pueden ser rápidamente tratadas actualizando el *Software* antivirus, requieren mayor diagnóstico antes de que se pueda diseñar un plan para erradicarlas.

Para poder decidir acerca de cómo proceder, debe saber con qué clase de virus está tratando y las posibles repercusiones en su entorno. Aprenda sobre cómo los virus se propagan, sus métodos de ataque y el posible daño que pueden provocar. No debe depender únicamente de la información que le suministran sobre los virus sus proveedores de antivirus o de detección de intrusos; aunque debería considerarlos como árbitros en caso de informaciones conflictivas.

Una vez que haya identificado el virus, debe hacer una buena evaluación concerniente a la gravedad de la infección. Formúlese las siguientes preguntas para evaluar el vector infeccioso y para identificar las maneras más rápidas de evitar que el virus continúe propagándose:

2.3.5. El Virus penetra en la red

Si penetra a través del correo electrónico SMTP¹², el filtrado de contenidos debería detenerlo. Si penetra porque se explora un servidor infectado, puede bloquear las direcciones URL¹³. Si llega a través de las aplicaciones entre iguales o de las

¹² SMTP: Simple Mail Transfer Protocol (SMTP) Protocolo Simple de Transferencia de correo

¹³ URL: Localizador de Recurso Uniforme (en inglés Uniform Resource Locator)

aplicaciones de chat en *Internet*, trate de bloquear los puertos específicos utilizados por estas aplicaciones en el *firewall*.

¿Está el virus consciente de la existencia de la red y se propaga a través de la utilización compartida de archivos? Si así es, es más probable que se propague a otros sistemas, a no ser que se tenga la costumbre de eliminar los archivos compartidos de administración y deshabilitar las funciones de la utilización compartida de archivos.

¿Utiliza el virus los programas grupales o los *gateways* del correo electrónico para propagarse más? Si se utiliza sistemas internos del correo electrónico para diseminarse, nuevamente el filtrado de contenidos debería detenerlo. Si, en cambio, se instala su propio servidor SMTP para enviar correos electrónicos, se podría bloquear temporalmente el puerto 25 en el *firewall*.

¿Penetra en el entorno a través de agujeros de seguridad? En ese caso, se debería aplicar parches de *Software* para detener su infiltración.

Igualmente, se debe evaluar la propagación de la infección. Revise todas las computadoras de administración de uso prioritario y de uso continuo y, si algunas se hallan infectadas, se debe desconectarlas de la red. Puesto que es más probable que otros sistemas hayan accedido a ellas, es probable que el virus se haya esparcido más. Por consiguiente, se debe considerar la desconexión de la red de los servidores no infectados, para prevenir su contagio.

Si el virus sabe de la existencia de la red, se debe revisar para comprobar si algunas cuentas del administrador de la red se encuentran comprometidas. Esto ocurre a menudo debido a la infección de un sistema del administrador de la red. Cuando sucede, es probable que el virus que sabe de la existencia de la red se haya propagado mucho. Seguramente, tendrá que suprimir las cuentas de usuario afectadas.

2.3.6. Esfuerzos

Si usted no lo ha hecho todavía, es hora de reunir a un equipo para enfrentarse a la amenaza, el cual será responsable de determinar las opciones, de recomendar la solución apropiada e implementar aquella que sea seleccionada. Utilice su equipo y procedimientos de respuesta a incidentes, si dispone de ellos. En caso contrario, necesita reunir un equipo.

En primer lugar, nombre al líder del equipo, quien deberá supervisar todo el proceso diseñado para la solución de la infección viral y luego tenga en cuenta los siguientes recursos para su equipo:

Personal de información, el cual a menudo lanza la primera voz de alarma, y continúa recibiendo las llamadas de los usuarios que se quejan de la aparición de los indicadores de infección. Ellos mismos pueden comunicar a los usuarios los procesos para responder al ataque.

Si el equipo de respuesta a incidentes existe, éste deberá ser el agente coordinador, investigando las especificaciones del virus, las soluciones recomendadas y los puntos iniciales de la infección.

Personal de estaciones de trabajo y de servidores operacionales, el cual podrá rastrear los sistemas infectados, identificar los puntos prioritarios que necesitan protección y comunicar la solución a los usuarios. Ambos equipos pueden requerir el parcheo de los sistemas como parte de dicho proceso de respuesta. Personal de redes, que puede tener que bloquear las conexiones en el perímetro y/o segmentar la red para contener la propagación del virus. Además, pueden revisar los registros de *firewall* y de los *routers* para localizar los puntos iniciales de la infección.

El equipo de mensajería, que puede ser llamado para poner fuera de servicio o para reconfigurar los servidores de correo electrónico, o para aplicar parches, como parte del proceso de respuesta. Puede también revisar los registros del correo electrónico en busca de indicadores de los puntos iniciales de infección. Puede necesitarse un representante legal, para prestar asistencia en las actividades de investigación forense, una vez que se haya contenido el virus.

Se puede requerir un contable, para determinar el impacto financiero del incidente viral y la conveniencia de llevar a cabo una investigación. El personal de Relaciones Públicas puede ser indispensable, si el virus o la respuesta afecta a los socios comerciales, a los clientes o a terceros. El Personal de Recursos Humanos puede involucrarse, si el incidente viral es el resultado de la violación de las políticas empresariales por parte de un empleado o de un contratista.

La participación de los altos directivos puede ser necesaria, si las decisiones que implique la respuesta al ataque viral llegasen, probablemente, a afectar negativamente a la compañía, a los socios, clientes o demás. Los representantes de las unidades y de los departamentos comerciales afectados, deben mantenerse informados.

Como parte del esfuerzo de recuperación, puede requerir asistencia externa. A medida que se comience a desarrollar el proceso de respuesta a incidentes, necesitará establecer la manera de comunicar el problema a todos los equipos, el modo de reportar el problema apropiadamente, la forma de hacerle un seguimiento al progreso de la solución y la oportunidad para involucrar a cada equipo en tal proceso. Una comunicación clara resulta invaluable para ayudar a determinar la causa del virus, los puntos de infección inicial, la propagación de la infección y para coordinar apropiadamente la respuesta.

2.3.7. Contención del ataque

No siempre basta con actualizar las definiciones antivirus, ni explorar los sistemas para eliminar el virus de su entorno. Debe localizar todos los sistemas que se encuentran infectados y limpiarlos todos completamente, para recuperar el control de su red. En muchos casos, es más fácil decirlo que hacerlo y puede demorarse. Al mismo tiempo, debe tomar medidas para contener la continua amenaza.

Si ésta existe, base sus medidas de contención en su política de respuesta a incidentes y ejecutarla solamente después de que haya reunido suficiente información para tomar una decisión apropiada y de que haya obtenido la aprobación gerencial necesaria en cada medida. El plan de contención debería incluir cuándo las medidas provisionales deben revertirse, para volver al funcionamiento normal.

Todas las actividades de contención deberían controlarse de forma centralizada, después de notificarlo a las partes afectadas. Esto podría incluir personas a cargo de los servidores de mensajería, de los servidores *Web*, del acceso a *Internet*, de los servidores de archivo y de impresión y/o de los servidores de aplicación. A veces es más fácil determinar quién está afectado, basándose en las divisiones por departamentos o en la localización de la red.

Si el virus puede propagarse a través del correo electrónico y el acceso a HTTP o FTP, concéntrese en actualizar primero la protección en sus servidores de correo electrónico o de *Software* de grupo, en los servidores *proxy* y en los servidores *gateway* de correo electrónico SMTP; estos son los medios a través de los cuales se propaga más rápidamente la infección, tanto interna como externamente. Actualizar la protección puede implicar la instalación de parches de *Software*, la actualización de las firmas de los virus, la implementación de filtrado de contenidos y otras medidas.

Si la infección se está propagando más rápidamente de lo que usted puede distribuir las reparaciones, deberá contener el ataque deshabilitando los servicios que permiten la llegada de peticiones. Esto es especialmente importante en las situaciones siguientes:

Cuando el proveedor no tiene todavía disponibles las firmas de antivirus.
Cuando no es posible el filtrado del contenido debido al contenido cambiante.
Cuando los usuarios no están adiestrados apropiadamente con respecto a las amenazas virales y al papel que deben desempeñar en cuanto a la protección.
Cuando estén disponibles, consulte las publicaciones técnicas del proveedor sobre el virus, siempre que se piense adelantar acciones de contención. Las posibles acciones incluyen:

Dividir la red usando *firewalls*, *routers* o conmutadores. Reconfigurar el servidor DNS¹⁴ para desactivar el correo electrónico SMTP que entra.

Cambiar el *Software* de filtrado de contenidos para bloquear todos los archivos anexos de correo electrónico. Cambie el *Software* de filtrado de contenidos para bloquear el correo electrónico que contenga algunas cadenas de texto. Bloquee los accesos HTTP o FTP a *Internet*. Considere el enviar a casa a los usuarios de los departamentos que hayan sido seriamente afectados por la interrupción o retire de la red las estaciones de trabajo que no hayan resultado afectadas para que los usuarios puedan seguir trabajando localmente. Considere el proteger contra escritura los datos importantes que sean accesibles al virus en sistemas de alta prioridad.

Crear archivos ficticios en sistemas que no hayan sido infectados, para evitar la infección del sistema. Si el riesgo es extremadamente alto, desconecte completamente la red de *Internet*. También debería usted pensar en desactivar los

¹⁴ DNS: El sistema de nombre de dominio (en inglés Domain Name System)

servicios que proveen peticiones que salen, con el fin de proteger a sus socios comerciales, a los clientes y a otras empresas. Esto puede incluir el desactivar los mensajes electrónicos que salen o algunos puertos en su *firewall*.

2.4. TIPOS DE DETECCIÓN

En primer lugar se clasifican según la manera en que detectan las intrusiones. Se categoriza las intrusiones en dos tipos principales, cuya distinción es importante porque conducirán a sistemas de detección esencialmente muy diferentes.

- Los usos indebidos son ataques bien definidos contra debilidades conocidas de los sistemas. Se los puede detectar buscando la ocurrencia de determinadas acciones concretas.
- Las anomalías se basan en la observación de desviaciones de los patrones de uso normales en el sistema. Se las detecta construyendo previamente un perfil del sistema a monitorizar y posteriormente estudiando las desviaciones que se produzcan con respecto a este perfil.

Las intrusiones por uso indebido siguen patrones bien definidos, por lo que se pueden detectar realizando búsqueda de patrones en el tráfico de red y en los ficheros de registro.

Las intrusiones por anomalía se detectan observando desviaciones significativas del comportamiento habitual. Para ello se mide una serie de parámetros (carga de CPU, número de conexiones de red en una unidad de tiempo, número de procesos, entre otros). Considerando que una intrusión involucrará un uso anormal del sistema, se pueden detectar las violaciones de seguridad a partir de patrones anormales de uso.

Los detectores de anomalías conocen, bien porque han sido programados por un experto, bien porque han pasado por una fase previa de aprendizaje, la actividad que resulta “normal” en el seno de un sistema. Mediante métodos estadísticos se intentará posteriormente comparar la información recibida en cada instante con el modelo de actividad válida, y aquello que se aparte excesivamente será etiquetado como intrusión. Esta comparación se puede realizar por técnicas estadísticas, por sistemas expertos basados en reglas, con redes neuronales, o con algún otro tipo de reconocimiento de patrones que pueda emitir con una certeza razonable si una determinada secuencia de eventos en un sistema forma parte del funcionamiento ordinario del mismo.

Es difícil detectar intrusiones por anomalías. No hay patrones fijos que se puedan monitorizar, por lo que se usan aproximaciones “borrosas” que suelen producir altas tasas de error. La correlación de los datos recibidos por los sensores es en la actualidad un área de investigación sujeta a estudio. Se persigue minimizar el número de falsos positivos (falsas alarmas) y de falsos negativos (ataques reales que pasan inadvertidos al sistema).

El paradigma de detección de anomalías parece bastante potente, pues en principio es capaz de detectar todo tipo de ataques, incluso ataques desconocidos hasta la fecha de su ocurrencia. En el caso de sistemas basados en reglas, exigen de un experto que pueda introducir correctamente dicho conjunto, que ha de ser periódicamente actualizado conforme las prácticas varíen. En el caso de sistemas basados en aprendizaje puede ocurrir que un atacante varíe muy lentamente su comportamiento para hacer casar una actividad maliciosa dentro de lo aceptable por el nuevo modelo aprendido. Los sistemas informáticos son por naturaleza muy cambiantes y los detectores de anomalías pueden producir una tasa de falsos positivos inaceptable.

En la práctica se han extendido más los detectores de usos indebidos, que se basan en una base de datos de ataques conocidos, con una serie de reglas o

“signaturas” que caracterizan los ataques y que permiten aseverar con prácticamente total certeza que se está intentando perpetrar un ataque. Estos sistemas solo pueden detectar fallos conocidos, para los que se haya introducido la signatura correspondiente en la lista. Dado que cada día aparecen nuevas vulnerabilidades, es importante que estos sistemas dispongan de mecanismos para actualizar frecuentemente la base de signaturas.

2.4.1. Fuentes de información Forense

Dependiendo de las fuentes de información que se utilicen, los sensores usados por los IDS¹⁵ se clasifican en dos tipos: de red y de máquina. Cada tipo tiene unas capacidades diferentes en cuanto a los eventos detectables, por lo que en la práctica los IDS suelen nutrirse de sensores de ambos tipos. En la terminología tradicional de IDS, se habla de NIDS (Sistemas de Detección de Intrusos de Red) y de HIDS (Sistemas de detección de intrusos de máquina). En los sistemas híbridos o distribuidos, que abarcan más de un solo nodo, se habla de sensores: un solo sistema de detección de intrusiones puede alimentarse de más de un sensor. Como la mayoría de los sistemas de IDS comerciales son aparatos independientes dotados de sensores de red que se conectan sin tener que instalar nada en ninguna otra máquina, se ha abusado bastante del término NIDS.

2.4.1.1. NIDS

Sistemas de detección de intrusos por red, estos sistemas disponen de una o varias interfaces de red conectadas a determinados puntos estratégicos de la red. Monitorizan el tráfico que pasa por dichos puntos en busca de tráfico malicioso. Aunque estos sistemas en principio son dispositivos absolutamente pasivos, con frecuencia se colocan los NIDS en cortafuegos y enrutadores, de manera que el

¹⁵ IDS: Un sistema de detección de intrusos ó Intrusion Detection System

propio sistema puede forzar el cierre de conexiones y modificar reglas de filtrado de una manera más directa.

Mediante uno solo de estos sistemas se puede monitorizar el tráfico tanto interno como externo de una red para muchas máquinas. Los NIDS no suelen controlar toda la red sino determinados puntos estratégicos. La mayoría de las redes hoy en día son conmutadas, así que colocar los sensores de red suele implicar utilizar conmutadores especiales con un puerto "monitor" que reproduce todo el tráfico recibido en cualquiera de los puertos.



Figura N°1: Seguridad y detección de Intrusos
Fuente: [AWE-2006]

Este tipo de sistemas son bastante rápidos de instalar y mantener, y no dependen del sistema operativo instalado en las máquinas cubiertas. Suelen ser invisibles para los atacantes, por lo que los registros de sucesos que almacenan son poco vulnerables a la eliminación o alteración maliciosa, y suponen un recurso valioso para el almacenamiento de pruebas.

Diferentes ubicaciones de los NIDS proporcionarán diferentes perspectivas de la seguridad de la red. Colocados fuera de la corta fuegos permiten evaluar los

ataques que se intentan producir aunque no alcancen a los servidores internos, mientras que si se colocan en el interior de los cortafuegos permiten evaluar si este está bien configurado.

2.4.1.2. HIDS

Sistemas de detección de intrusos de máquina. Así como los NIDS se instalan en determinados puntos de la infraestructura de red, los HIDS se instalan en las máquinas que componen la red: tanto servidores como estaciones de trabajo. Un sensor, instalado directamente como un módulo sobre una máquina, dispone de información de mayor nivel semántico que los NIDS: llamadas al sistema, eventos complejos dentro de aplicaciones de alto nivel, etc. Un sistema basado únicamente en red tendría que ser mucho más complejo para “entender” la gran diversidad de protocolos que existen, y los que se implementan por encima de éstos. Por otra parte, la tendencia actual al uso de conexiones encriptados, de indiscutible interés para mejorar la seguridad de los sistemas, hace que un sistema que solo escuche la red disponga de muy poca información para distinguir el tráfico malicioso del aceptable.

El tráfico en una conexión SSH¹⁶ o SSL¹⁷ es absolutamente inaccesible a un NIDS, aunque en el caso de SSL se han desarrollado cortafuegos que interceptan las conexiones, realizando una especie de ataque “hombre en el medio” que le permite analizar el contenido de conexiones que de otra manera sería inaccesible.

Los HIDS tienen acceso a los archivos de registro de lo que realmente sucedió, por lo que pueden conocer de manera fiable si un ataque fue exitoso o no, información generalmente no disponible para los NIDS. Un sensor de máquina dispone de información específica del sistema y las aplicaciones, como inicios/cierres de sesión, acceso a ficheros, llamadas al sistema (pueden utilizarlas para saber el

¹⁶ SSH: protocolo informático que sirve para acceder a máquinas remotas

¹⁷ SSL: protocolo informático que sirve para acceder a máquinas remotas de tipo local

disco libre, la ocupación de la red, etc), y otros eventos, incluyendo aquellos que se originan localmente sin generar tráfico de red.

Tienen sobre los NIDS la ventaja de que permiten acceder a la información que por la red transita encriptado y que por lo tanto es opaca a ellos. Periódicos o de tiempo real Así como los NIDS suelen dar respuesta en tiempo real, los primeros HIDS se ejecutaban periódicamente para buscar indicios de intrusión. Después se fue reduciendo el intervalo entre la ocurrencia del evento y su análisis, hasta el punto que es posible gestionar los eventos en el instante de su registro. Los sistemas de red implementados como parte de la pila de red de las máquinas protegidas ofrecen las mismas prestaciones de respuesta inmediata que los NIDS.

Activos o pasivos: Los primeros IDS eran pasivos, se limitaban a informar de los intentos de intrusión al administrador. De poco sirve detectar un ataque para que horas después el administrador reciba un mensaje que informe de que se vio la intrusión pero no se intentó hacer nada por abortarla. Los IDS activos son capaces de tomar acciones correctivas orientadas a detener ataques en el mismo instante en que se producen.

Centralizados o distribuidos: Cuando la red de una organización adquiere una envergadura determinada, ya no es factible analizar todo el tráfico en un sólo punto sin producir una degradación del rendimiento. En tal caso se instalan sistemas distribuidos, que disponen de varios sensores repartidos por diversas máquinas y puntos de la red, que se comunican con un nodo central donde se reciben todas las informaciones relevantes y donde se cruzan los datos para disponer de una visión más amplia del sistema como conjunto y detectar con mayor fiabilidad eventuales ataques. Esto permite producir además una única respuesta a intrusiones visibles desde varios puntos de la red.

Evasión de IDS: Es posible que el sistema no sea capaz de detectar una determinada instancia de ataque conocido al ser incapaz de encontrar la

coincidencia con el patrón de búsqueda, si el atacante se las arregla para introducir pequeñas variaciones en su interacción con la máquina precisamente con el objetivo de evadir el IDS. Por ejemplo, algunas estrategias de evasión explotan leves diferencias en la manera en que la pila TCP reensambla fragmentos, o la manera en que se procesan paquetes inválidos, etcétera. La mayoría de los productos IDS de hoy en día incluyen protecciones contra las técnicas de evasión de IDS.

Sistemas de decepción: Son un tipo especial de sistema de detección de intrusiones orientadas a atraer la atención de potenciales intrusos para que no ataquen a los sistemas reales y para obtener información acerca de sus métodos. Son los llamados *honeypots* (tarros de miel): máquinas simuladas, verosímiles y relativamente poco ocultas. Dado que ningún usuario legítimo debería querer jamás intentar conectarse a un *honeypot*, toda conexión al mismo puede informarse inmediatamente y etiquetarse como un intento de intrusión.

Los *honeypots* están configurados para registrar los eventos extensamente. La irrupción de un intruso en estas máquinas permite a los administradores obtener información sobre su *modus operandi*, e incluso recabar pruebas o indicios que pudieran inculpar al delincuente en un juicio.

Análisis forense: Los IDS ofrecen un interesante servicio para el análisis forense después de la consumación de ataques. Es posible que un IDS no haya sido capaz de detener la acción de un atacante, pero sí puede haber guardado un registro de los mensajes que transitaron por la red a tal efecto. Aunque cualquier atacante que tenga cierto nivel hará todo lo posible por borrar sus huellas, falsificar direcciones, explotar máquinas de terceros para enmascararse, etcétera, toda información que se almacene puede ayudar a seguir la pista del atacante, a mejorar los sistemas de detección y reacción automatizada a dichos ataques, e incluso como indicios ante instancias judiciales.

Algunas herramientas disponibles *Snort*, uno de los sistemas más utilizados actualmente, es un sistema de código abierto de detección de intrusiones de red, capaz de llevar a cabo análisis de tráfico en tiempo real y registro de paquetes en redes IP. Puede efectuar análisis de protocolos, búsqueda de cadenas o patrones en el contenido y puede utilizarse para detectar una gran variedad de ataques y sondeos, tal como desbordamientos de búfer, *escaneos* invisibles, ataques CGI¹⁸, sondeos SMB¹⁹, intentos de determinación del sistema operativo, y otros.

Snort utiliza un flexible lenguaje de reglas para describir el tráfico que debería recoger o pasar, así como un motor de detección que hace uso de una arquitectura de *plugins* modular. Entre su base de reglas incluye miles de comprobaciones en busca de ataques de denegaciones de servicio. Ofrece la posibilidad de alertar en tiempo real, al incorporar mecanismos para registrar a *syslog*, a fichero, a sockets *Unix*, o mediante Samba, enviar mensajes emergentes a clientes *Windows*. Además de ser un sistema completo de detección de intrusiones de red, sirve como analizador de paquetes al estilo de *tcpdump*, y como herramienta para registrar el tráfico. Se puede compilar en una veintena de plataformas distintas, tanto sistemas *Unix* como Win32.

Prelude: *Snort* es el IDS de red libre más potente, pero en su arquitectura no contempla la posibilidad de usar sensores de máquina, lo cual motivó la aparición del proyecto, también libre, *Prelude*, que utiliza una arquitectura distribuida, con canales autenticados y encriptados, y sensores para diversos sistemas operativos. *Prelude* no pretende reinventar la rueda en IDS de red, y de hecho es capaz de nutrirse de *Snort*, e incluso incluye él mismo un motor que utiliza los ficheros de reglas de su predecesor.

¹⁸ CGI: Common Gateway Interface, una tecnología que se usa en los servidores web

¹⁹ SMB; Server Message Block o SMB es un Protocolo de red

Intrudec: El ITI²⁰ está desarrollando un prototipo de sistema de detección de intrusiones, *Intrudec*. Se trata de una arquitectura distribuida, con tolerancia a fallos, altamente modular, con soporte para sensores, tanto de red como de máquina, que se comunican de manera segura para permitir la correlación de los diversos eventos ocurridos en distintos puntos de la red y en los distintos sistemas monitorizados. Para la correlación se utilizan diversos algoritmos, cuyo desarrollo y ajuste constituyen la labor de investigación principal del grupo de Sistemas Fiables en el área de la detección de intrusiones. *Intrudec* complementa al proyecto *TigerWeb*, que este grupo ha estado desarrollando y manteniendo en los últimos dos años. *TigerWeb* es un sistema de detección remota de vulnerabilidades accesible vía *Web* que proporciona informes bien organizados y en castellano, orientados a ser entendidos por personal no experto en el área de la seguridad de los sistemas informático.

2.5. CÓMPUTO FORENSE

El cómputo forense, también llamado informática forense, computación forense, análisis forense digital o exanimación forense digital es la aplicación de técnicas científicas y analíticas especializadas a infraestructura tecnológica que permiten identificar, preservar, analizar y presentar datos que sean válidos dentro de un proceso legal.

Dichas técnicas incluyen reconstruir el bien informático, examinar datos residuales, autenticar datos y explicar las características técnicas del uso aplicado a los datos y bienes informáticos.

Como la definición anterior lo indica, esta disciplina hace uso no sólo de tecnología de punta para poder mantener la integridad de los datos y del procesamiento de los mismos; sino que también requiere de una especialización y conocimientos

²⁰ ITI: Intrudec Technology information

avanzados en materia de informática y sistemas para poder detectar dentro de cualquier dispositivo electrónico lo que ha sucedido.

La importancia de éstos y el poder mantener su integridad se basa en que la evidencia digital o electrónica es sumamente frágil. El simple hecho de darle doble clic a un archivo modificaría la última fecha de acceso del mismo.

Adicionalmente, un examinador forense digital, dentro del proceso del cómputo forense puede llegar a recuperar información que haya sido borrada desde el sistema operativo.

2.5.1. Dispositivos a Analizar

La infraestructura informática que puede ser analizada puede ser toda aquella que tenga una Memoria (informática), por lo que se pueden analizar los siguientes dispositivos:

- Disco duro de una Computadora o Servidor
- Documentación referida del caso.
- Logs de seguridad.
- Credenciales de autenticación
- Trazo de paquetes dentro de redes.
- Teléfono Móvil o Celular, parte de la telefonía celular,
- Agendas Electrónicas
- Dispositivos de GPS²¹.
- Impresora
- Memoria USB²²

²¹ GPN: El Global Positioning System (Sistema de Posicionamiento Global)

²² USB: El Universal Serial Bus (bus universal en serie)

2.5.2. Pasos del cómputo forense

El proceso de análisis forense a una computadora se describe a continuación:

2.5.2.1. Identificación

Es muy importante conocer los antecedentes, situación actual y el proceso que se quiere seguir para poder tomar la mejor decisión con respecto a las búsquedas y la estrategia de investigación. Incluye muchas veces la identificación del bien informático, su uso dentro de la red, el inicio de la cadena de custodia (proceso que verifica la integridad y manejo adecuado de la evidencia), la revisión del entorno legal que protege el bien y del apoyo para la toma de decisión con respecto al siguiente paso una vez revisados los resultados.

2.5.2.2. Preservación

Este paso incluye la revisión y generación de las imágenes forenses de la evidencia para poder realizar el análisis. Dicha duplicación se realiza utilizando tecnología de punta para mantener la integridad de la evidencia y la cadena de custodia que se requiere. Al realizar una imagen forense, se hace referencia al proceso que se requiere para generar una copia "bit-a-bit" de todo el disco, el cual permitirá recuperar en el siguiente paso, toda la información contenida y borrada del disco duro. Para evitar la contaminación del disco duro, normalmente se ocupan bloqueadores de escritura de *Hardware*, los cuales evitan el contacto de lectura con el disco, lo que provocaría una alteración no deseada en los medios.

2.5.2.3. Análisis

Proceso de aplicar técnicas científicas y analíticas a los medios duplicados por medio del proceso forense para encontrar pruebas de ciertas conductas. Se pueden realizar búsquedas de cadenas de caracteres, acciones específicas del o de los

usuarios de la máquina como son el uso de dispositivos de USB (marca, modelo), búsqueda de archivos específicos, recuperación e identificación de correos electrónicos, recuperación de los últimos sitios visitados, recuperación del caché del navegador de *Internet*, etc.

Se explica de la forma más sencilla posible el uso de algunas herramientas que puede facilitar la tarea a la hora de realizar un análisis forense en entornos *Windows*.

Cuando un usuario no autorizado toma el control de un sistema, éste puede instalar múltiples *backdoors* (puertas traseras) que le permitan entrar al sistema en un futuro, aunque parchemos la vulnerabilidad original. Se denomina análisis forense al proceso de analizar una copia completa de un sistema que ha sufrido una intrusión o ataque.

El análisis forense permite obtener la mayor cantidad posible de información sobre:

- El método utilizado por el atacante para introducirse en el sistema
- Las actividades ilícitas realizadas por el intruso en el sistema
- El alcance y las implicaciones de dichas actividades
- Las “puertas traseras” instaladas por el intruso

Realizando un análisis forense permitirá, se puede una persona recuperar de un incidente de una manera más segura y evitando en la medida de lo posible que se repita la misma situación en cualquiera de las máquinas.

Un buen análisis forense debe dar respuestas a varias cuestiones, entre las que se encuentran las siguientes:

- ¿En qué fecha exacta se ha realizado la intrusión o cambio?
- ¿Quién realizó la intrusión?

- ¿Cómo entró en el sistema?
- ¿Qué daños ha producido en el sistema?

Si una vez realizado el análisis forense no se conoce con exactitud las respuestas a estas preguntas, no se tendrá un análisis funcional. Esto puede derivar en futuros ataques, bien por la misma persona, o bien por diferentes medios de intrusión que se desconozca.

2.5.2.4. Presentación

Es el recopilar toda la información que se obtuvo a partir del análisis para realizar el reporte y la presentación a los abogados, la generación (si es el caso) de una pericial y de su correcta interpretación sin hacer uso de tecnicismos.

2.5.3. Finalidad de la Informática Forense

La Informática forense permite la solución de conflictos tecnológicos relacionados con seguridad informática y protección de datos. Gracias a ella, las empresas obtienen una respuesta a problemas de privacidad, competencia desleal, fraude, robo de información confidencial y/o espionaje industrial surgidos a través de uso indebido de las tecnologías de la información. Mediante sus procedimientos se identifican, aseguran, extraen, analizan y presentan pruebas generadas y guardadas electrónicamente para que puedan ser aceptadas en un proceso legal.

- ¿Para qué sirve? Para garantizar la efectividad de las políticas de seguridad y la protección, tanto de la información como de las tecnologías que facilitan la gestión de esa información.
- ¿En qué consiste? Consiste en la investigación de los sistemas de información con el fin de detectar evidencias de la vulneración de los sistemas.

- ¿Cuál es su finalidad? Cuando una empresa contrata servicios de Informática forense puede perseguir objetivos preventivos, anticipándose al posible problema u objetivos correctivos, para una solución favorable una vez que la vulneración y las infracciones ya se han producido.
- ¿Qué metodologías utiliza la Informática forense? Las distintas metodologías forenses incluyen la recogida segura de datos de diferentes medios digitales y evidencias digitales, sin alterar los datos de origen. Cada fuente de información se cataloga preparándola para su posterior análisis y se documenta cada prueba aportada. Las evidencias digitales recabadas permiten elaborar un dictamen claro, conciso, fundamentado y con justificación de las hipótesis que en él se barajan a partir de las pruebas recogidas.
- ¿Cuál es la forma correcta de proceder? Y, ¿por qué? Todo el procedimiento debe hacerse teniendo en cuenta los requerimientos legales para no vulnerar en ningún momento los derechos de terceros que puedan verse afectados. Ello para que, llegado el caso, las evidencias sean aceptadas por los tribunales y puedan constituir un elemento de prueba fundamental, si se plantea un litigio, para alcanzar un resultado favorable.

2.5.4. Objetivos de la informática Forense

La utilización de la informática forense con una finalidad preventiva, en primer término. Como medida preventiva sirve a las empresas para auditar, mediante la práctica de diversas pruebas técnicas, que los mecanismos de protección instalados y las condiciones de seguridad aplicadas a los sistemas de información son suficientes. Asimismo, permite detectar las vulnerabilidades de seguridad con el fin de corregirlas. Cuestión que pasa por redactar y elaborar las oportunas

políticas sobre uso de los sistemas de información facilitados a los empleados para no atentar contra el derecho a la intimidad de esas personas.

Por otro lado, cuando la seguridad de la empresa ya ha sido vulnerada, la informática forense permite recoger rastros probatorios para averiguar, siguiendo las evidencias electrónicas, el origen del ataque (si es una vulneración externa de la seguridad) o las posibles alteraciones, manipulaciones, fugas o destrucciones de datos a nivel interno de la empresa para determinar las actividades realizadas desde uno o varios equipos concretos.

2.6. TÉCNICAS ANTI FORENSE

Uno de los grandes problemas a los que se enfrenta a la hora de realizar el análisis forense, es la detección de aplicaciones anti forense que han podido ser utilizadas para ocultar o eliminar información que pudiera estar en un sistema. El gran inconveniente de detectar este tipo de herramientas en un equipo, es la inquietud de saber que este equipo, portaba información de interés pero que desgraciadamente no se podrá tener acceso a ella.

Este tipo de herramientas sigue mecanismos diferentes todas tienen persiguen un mismo objetivo, dificultar la trazabilidad en un escenario forense. Podemos diferenciarlas en los siguientes tipos:

- Las enfocadas a la eliminación de la información.
- Las enfocadas a la ofuscación de la información.
- Las enfocadas a generar la incertidumbre en la investigación.

Todas por ellas mismas son ciertamente peculiares, tanto en su uso como en los fines que persiguen. El buen uso (o mal uso según la circunstancia) permiten que un potencial delito pudiera quedar impune. Desgraciadamente la falta de evidencias (objetivo que persiguen este tipo de aplicaciones), limita la posibilidad de

enjuiciamiento. La “clara evidencia” del empleo de una herramienta de anti forense, no permite determinar a efectos jurídicos el posible hecho delictivo, debido a la ausencia “clara de evidencias”, a lo sumo la existencia de conjeturas. No obstante y afortunadamente no siempre el empleo de estas herramientas es definitivo, puesto que aunque se elimina información relevante, no se hace realmente extensivo a los ficheros de carácter temporal que pudiera estar siendo utilizado por el Sistema Operativo y que puede quedar revelado en un análisis forense.

En análisis forense realizado con la herramienta FTK²³ *Access data* en un entorno de laboratorio, emulando las posibles acciones en un potencial delito, revelaba el uso de la herramienta *Evidence Eliminator*, mediante el empleo de estas firmas. Esta aplicación está basada en un interface cómodo, que permite realizar varias pasadas de 1 y/o 0 (*Wipe*) para sobrescribir determinada información que pudiera ser comprometida. Puesto que el tiempo que se tarda en realizar estas operaciones puede ser mucho y se pueden limitar las opciones de eliminación. En esta circunstancias en el análisis en el fichero de paginación, se reveló (puesto que no fue eliminado debido al tiempo que tardaba en realizarse la operación) las acciones realizadas en el equipo y la recuperación de información que pudiera ser transcendental al caso.

La existencia o la detección del uso de herramientas anti forense, debieran incitar a una investigación, con más ahínco si cabe, puesto que denota a través de su uso la posible importancia de las evidencias. El problema al que se enfrenta el investigador, es la desventaja moral con la que parte al saber que en un porcentaje muy alto de las circunstancias, sus esfuerzos no obtendrán ningún fruto.

Dentro de los tipos de herramientas que se han comentado previamente, podríamos citar algunas que identifican claramente el objetivo a perseguir.

²³ FTK: Final Turn Kill

- Eliminación de evidencias. Además de la citada *evidence eliminator*, se podría incluir en este grupo la *Suite Dban*. Esta herramienta en sí, es un disco de arranque que permite la eliminación segura de un disco. Además del posible uso delictivo que se pueda dar, puede ser también utilizado como mecanismo de destrucción documental o eliminación segura de información crítica. Por ejemplo cuando un equipo de una organización vaya a ser desechado (el formateo del disco, únicamente no es una buena práctica).
- Dentro de los sistemas de ofuscación, herramientas de cifrado como *Truecrypt* o el empleo de técnicas de *esteganografía* o *malware*, permiten que alguien conocedor del hecho pueda acceder a la información dificultando no obstante una posible investigación.
- Las enfocadas a generar la incertidumbre, no tienen como objetivo la eliminación u ocultación de la información. Simplemente confundir con la información existente. Se tiene un ejemplo con el uso de la herramienta *TimeStomp*. Perteneciente al grupo de herramientas *MAFIA (Metasploit Anti-Forensic Investigation Arsenal)* del proyecto *Metasploit*, persigue alterar la información de tiempos de ficheros. Permite poner fechas inverosímiles, complicando el análisis al conseguir la ruptura de la línea temporal de la investigación.

Uno no sabe si es mejor detectar su empleo en un equipo analizado o no, porque a veces saber que se han empleado, dejan con la incertidumbre de no saber, que se ha podido hacer con ellas.

2.7.SEGURIDAD INFORMÁTICA

La valiosa información que una entidad tiene implica una seguridad informática óptima para asegurar que los recursos del sistema de información (material

informático o programas) de una organización sean utilizados de la manera que se decidió y qué el acceso a la información allí contenida, así como su modificación, sólo sea posible a las personas que se encuentren acreditadas y dentro de los límites de su autorización.

Se puede entender como seguridad un estado de cualquier tipo de información (informático o no) que indique que ese sistema está libre de peligro, daño o riesgo. Se entiende como peligro o daño todo aquello que pueda afectar su funcionamiento directo o los resultados que se obtienen del mismo. Para la mayoría de los expertos el concepto de seguridad en la informática es utópico porque no existe un sistema 100% seguro. Para que un sistema se pueda definir como seguro debe tener estas cuatro características:

- Integridad: La información sólo puede ser modificada por quien está autorizado y de manera controlada.
- Confidencialidad: La información sólo debe ser elegible para los autorizados.
- Disponibilidad: Debe estar disponible cuando se necesita.
- Irrefutabilidad (No repudio): El uso y/o modificación de la información por parte de un usuario debe ser irrefutable, es decir, que el usuario no puede negar dicha acción.

Dependiendo de las fuentes de amenaza, la seguridad puede dividirse en tres partes: seguridad física, seguridad ambiental y seguridad lógica.

En estos momentos la seguridad informática es un tema de dominio obligado por cualquier usuario de la *Internet*, para no permitir que su información sea comprometida.

Términos relacionados con la seguridad informática

- Activo: recurso del sistema de información o relacionado con éste, necesario para que la organización funcione correctamente y alcance los objetivos propuestos.
- Amenaza: es un evento que puede desencadenar un incidente en la organización, produciendo daños materiales o pérdidas inmateriales en sus activos.
- Impacto: medir la consecuencia al materializarse una amenaza.
- Riesgo: posibilidad de que se produzca un impacto determinado en un Activo, en un Dominio o en toda la Organización.
- Vulnerabilidad: posibilidad de ocurrencia de la materialización de una amenaza sobre un Activo.
- Ataque: evento, exitoso o no, que atenta sobre el buen funcionamiento del sistema.
- Desastre o Contingencia: interrupción de la capacidad de acceso a información y procesamiento de la misma a través de computadoras necesarias para la operación normal de un negocio.

Aunque a simple vista se puede entender que un Riesgo y una Vulnerabilidad se podrían englobar un mismo concepto, una definición más informal denota la diferencia entre riesgo y vulnerabilidad, de modo que se debe la vulnerabilidad está ligada a una Amenaza y el Riesgo a un Impacto.

La información (datos) se verá afectada por muchos factores, incidiendo básicamente en los aspectos de confidencialidad, integridad y disponibilidad de la misma. Desde el punto de vista de la empresa, uno de los problemas más importantes puede ser el que está relacionado con el delito o crimen informático, por factores externos e internos. Una persona no autorizada podría: Clasificar y desclasificar los datos, Filtrar información, Alterar la información, Borrar la información, Usurpar datos, Hojear información clasificada.

La medida más eficiente para la protección de los datos es determinar una buena política de copias de seguridad o *backups*: Copia de seguridad completa, Todos los datos (la primera vez), Copias de seguridad incrementales, Sólo se copian los ficheros creados o modificados desde el último *backup*, Elaboración de un plan de *backup* en función del volumen de información generada:

- Tipo de copias, ciclo de esta operación, etiquetado correcto.
- Diarias, semanales, mensuales: creación de tablas periódicamente hablando, solicitando un dinero como ya se dijo diarios, mensuales o semanales.

2.7.1. Objetivos

Los activos son los elementos que la seguridad informática tiene como objetivo proteger. Son tres elementos que conforman los activos:

- Información: Es el objeto de mayor valor para una organización, el objetivo es el resguardo de la información, independientemente del lugar en donde se encuentre registrada, en algún medio electrónico o físico.
- Equipos que la soportan: *Software*, *Hardware* y organización.

- Usuarios: Individuos que utilizan la estructura tecnológica y de comunicaciones que manejan la información.

2.7.2. Análisis de Riesgos

El activo más importante que se posee es la información y, por lo tanto, deben existir técnicas que la aseguren, más allá de la seguridad física que se establezca sobre los equipos en los cuales se almacena. Estas técnicas las brinda la seguridad lógica que consiste en la aplicación de barreras y procedimientos que resguardan el acceso a los datos y sólo permiten acceder a ellos a las personas autorizadas para hacerlo.

Existe un viejo dicho en la seguridad informática que dicta: "lo que no está permitido debe estar prohibido" y ésta debe ser la meta perseguida.

Los medios para conseguirlo son:

- Restringir el acceso (de personas de la organización y de las que no lo son) a los programas y archivos.
- Asegurar que los operadores puedan trabajar pero que no puedan modificar los programas ni los archivos que no correspondan (sin una supervisión minuciosa).
- Asegurar que se utilicen los datos, archivos y programas correctos en el procedimiento elegido.
- Asegurar que la información transmitida sea la misma que reciba el destinatario al cual se ha enviado y que no le llegue a otro.

- Asegurar que existan sistemas y pasos de emergencia alternativos de transmisión entre diferentes puntos.
- Organizar a cada uno de los empleados por jerarquía informática, con claves distintas y permisos bien establecidos, en todos y cada uno de los sistemas o aplicaciones empleadas.
- Actualizar constantemente las contraseñas de accesos a los sistemas de cómputo.

2.7.3. Elementos de un Análisis

Cuando se pretende diseñar una técnica para implementar un análisis de riesgo informático se pueden tomar los siguientes puntos como referencia a seguir:

- Construir un perfil de las amenazas que esté basado en los activos de la organización.
- Identificación de los activos de la organización.
- Identificar las amenazas de cada uno de los activos listados.
- Conocer las prácticas actuales de seguridad
- Identificar las vulnerabilidades de la organización.
 - Recursos humanos
 - Recursos técnicos
 - Recursos financieros
- Identificar los requerimientos de seguridad de la organización.

- Identificación de las vulnerabilidades dentro de la infraestructura tecnológica.
- Detección de los componentes claves
- Desarrollar planes y estrategias de seguridad que contengan los siguientes puntos:
 - Riesgo para los activos críticos
 - Medidas de riesgos
 - Estrategias de protección
 - Planes para reducir los riesgos.

2.7.4. Análisis de Impacto al Negocio

El reto es asignar estratégicamente los recursos para equipo de seguridad y bienes que intervengan, basándose en el impacto potencial para el negocio, respecto a los diversos incidentes que se deben resolver. Para determinar el establecimiento de prioridades, el sistema de gestión de incidentes necesita saber el valor de los sistemas de información que pueden ser potencialmente afectados por incidentes de seguridad.

Esto puede implicar que alguien dentro de la organización asigne un valor monetario a cada equipo y un archivo en la red o asignar un valor relativo a cada sistema y la información sobre ella. Dentro de los Valores para el sistema se pueden distinguir: Confidencialidad de la información, la Integridad (aplicaciones e información) y finalmente la Disponibilidad del sistema. Cada uno de estos valores es un sistema independiente del negocio, supongamos el siguiente ejemplo, un servidor *Web* público pueden poseer los requisitos de confidencialidad de baja (ya que toda la información es pública), pero de alta disponibilidad y los requisitos de integridad. En contraste, un sistema de planificación de recursos empresariales,

sistema puede poseer alta puntaje en los tres variables. Los incidentes individuales pueden variar ampliamente en términos de alcance e importancia.

2.7.5. Puesta en Marcha de una Política de Seguridad

Actualmente las legislaciones nacionales de los Estados, obligan a las empresas, instituciones públicas a implantar una política de seguridad. Generalmente se ocupa exclusivamente a asegurar los derechos de acceso a los datos y recursos con las herramientas de control y mecanismos de identificación. Estos mecanismos permiten saber que los operadores tienen sólo los permisos que se les dio.

La seguridad informática debe ser estudiada para que no impida el trabajo de los operadores en lo que les es necesario y que puedan utilizar el sistema informático con toda confianza. Por eso en lo referente a elaborar una política de seguridad, conviene:

- Elaborar reglas y procedimientos para cada servicio de la organización.
- Definir las acciones a emprender y elegir las personas a contactar en caso de detectar una posible intrusión
- Sensibilizar a los operadores con los problemas ligados con la seguridad de los sistemas informáticos.

Los derechos de acceso de los operadores deben ser definidos por los responsables jerárquicos y no por los administradores informáticos, los cuales tienen que conseguir que los recursos y derechos de acceso sean coherentes con la política de seguridad definida. Además, como el administrador suele ser el único en conocer perfectamente el sistema, tiene que derivar a la directiva cualquier problema e información relevante sobre la seguridad, y eventualmente aconsejar estrategias a poner en marcha, así como ser el punto de entrada de la

comunicación a los trabajadores sobre problemas y recomendaciones en término de seguridad informática.

2.7.6. Las Amenazas

Una vez que la programación y el funcionamiento de un dispositivo de almacenamiento (o transmisión) de la información se consideran seguras, todavía deben ser tenidos en cuenta las circunstancias "no informáticas" que pueden afectar a los datos, las cuales son a menudo imprevisibles o inevitables, de modo que la única protección posible es la redundancia (en el caso de los datos) y la descentralización.

Estos fenómenos pueden ser causados por:

- El usuario: causa del mayor problema ligado a la seguridad de un sistema informático (porque no le importa, no se da cuenta o a propósito).
- Programas maliciosos: programas destinados a perjudicar o a hacer un uso ilícito de los recursos del sistema. Es instalado (por inatención o maldad) en el ordenador abriendo una puerta a intrusos o bien modificando los datos. Estos programas pueden ser un virus informático, un gusano informático, un troyano, una bomba lógica o un programa espía o *Spyware*.
- Un intruso: persona que consigue acceder a los datos o programas de los cuales no tiene acceso permitido (*cracker*, *script kiddie* o *Script boy*, *viruxer*, etc.).
- Un siniestro (robo, incendio, por agua): una mala manipulación o una mal intención derivan a la pérdida del material o de los archivos.

- El personal interno de Sistemas. Las pujas de poder que llevan a disociaciones entre los sectores y soluciones incompatibles para la seguridad informática.

2.7.7. Tipos de amenazas

El hecho de conectar una red a un entorno externo nos da la posibilidad de que algún atacante pueda entrar en ella, con esto, se puede hacer robo de información o alterar el funcionamiento de la red. Sin embargo el hecho de que la red no sea conectada a un entorno externo no nos garantiza la seguridad de la misma. De acuerdo con el *Computer Security Institute* (CSI) de San Francisco aproximadamente entre 60 y 80 por ciento de los incidentes de red son causados desde adentro de la misma. Basado en esto podemos decir que existen 2 tipos de amenazas:

- Amenazas internas: Generalmente estas amenazas pueden ser más serias que las externas por varias razones como son:
 - Los usuarios conocen la red y saben cómo es su funcionamiento.
 - Tienen algún nivel de acceso a la red por las mismas necesidades de su trabajo.
 - Los IP y *Firewall* son mecanismos no efectivos en amenazas internas.
- Amenazas externas: Son aquellas amenazas que se originan de afuera de la red. Al no tener información certera de la red, un atacante tiene que realizar ciertos pasos para poder conocer qué es lo que hay en ella y buscar la manera de atacarla. La ventaja que se tiene en este caso es que el administrador de la red puede prevenir una buena parte de los ataques externos.

2.7.8. La amenaza Informática del futuro

Si en un momento el objetivo de los ataques fue cambiar las plataformas tecnológicas ahora las tendencias cibercriminales indican que la nueva modalidad es manipular los significados de la información digital. El área semántica, era reservada para los humanos, se convirtió ahora en el núcleo de los ataques debido a la evolución de la *Web* y las redes sociales, factores que llevaron al nacimiento de la nueva generación.

- Se puede afirmar que la *Web* de nueva generación otorga contenidos y significados de manera tal que pueden ser comprendidos por las computadoras, las cuales -por medio de técnicas de inteligencia artificial- son capaces de emular y mejorar la obtención de conocimiento, hasta el momento reservada a las personas”.
- Es decir, se trata de dotar de significado a las páginas *Web*, y de ahí el nombre de *Web* semántica o Sociedad del Conocimiento, como evolución de la ya pasada Sociedad de la Información

En este sentido, las amenazas informáticas que viene en el futuro ya no son con la inclusión de troyanos en los sistemas o *Software* espías, sino con el hecho de que los ataques se han profesionalizado y manipulan el significado del contenido virtual.

- La *Web* de la nueva generación basada en conceptos como elaborar, compartir y significar, está representando un desafío para los *hackers* que ya no utilizan las plataformas convencionales de ataque, sino que optan por modificar los significados del contenido digital, provocando así la confusión lógica del usuario y permitiendo de este modo la intrusión en los sistemas”, La amenaza ya no solicita la clave de *homebanking* del desprevenido usuario, sino que directamente modifica el balance de la

cuenta, asustando al internauta y, a partir de allí, sí efectuar el robo del capital”.

Para no ser presa de esta nueva ola de ataques más sutiles, se recomienda;

- Mantener las soluciones activadas y actualizadas.
- Evitar realizar operaciones comerciales en computadoras de uso público.
- Verificar los archivos adjuntos de mensajes sospechosos y evitar su descarga en caso de duda.

2.7.8.1. En el futuro

La incorporación de las denominadas "redes inteligentes" podría dificultar considerablemente las actividades de los Hackers.

- El Instituto Tecnológico de Georgia, EEUU, trabaja en un proyecto de desarrollo de redes neurológicas, que probablemente aumentarán la seguridad del tráfico digital.
- El nombre "red neurológica" se basa en las neuronas del cerebro humano, que aprenden de la experiencia, creando conexiones entre las distintas áreas del cerebro. Con todo, cabe precisar que no se trata de redes que estén en condiciones de pensar, sino de sistemas capaces de identificar patrones en el flujo digital y aprender de los intentos de intrusión.
- Hoy en día, los administradores de sistemas deben actualizar manualmente los sistemas de protección de las redes contra las embestidas de los sagaces piratas informáticos. Con la incorporación de redes inteligentes se hará más previsible y fácil la contención de los intrusos.

- Según Cannady, tales redes estarán incluso en condiciones de detectar máquinas que monitorizan ilegalmente el tráfico de la red para captar y apoderarse de información tal como números de tarjetas de crédito, contraseñas y otros datos confidenciales. La novedad es que las redes neurológicas detectarán ese tipo de máquinas sin que sus operadores se percaten.

2.7.9. Técnicas para asegurar el sistema

- Codificar la información: Criptología, Criptografía y Criptociencia, contraseñas difíciles de averiguar a partir de datos personales del individuo.
- Vigilancia de red: Zona desmilitarizada.
- Tecnologías repelentes o protectoras: cortafuegos, sistema de detección de intrusos - *antispyware*, antivirus, llaves para protección de *Software*, etc. Mantener los sistemas de información con las actualizaciones que más impacten en la seguridad.

2.7.9.1. Consideraciones de *Software*

Tener instalado en la máquina únicamente el *Software* necesario reduce riesgos. Así mismo tener controlado el *Software* asegura la calidad de la procedencia del mismo (el *Software* obtenido de forma ilegal o sin garantías aumenta los riesgos). En todo caso un inventario de *Software* proporciona un método correcto de asegurar la reinstalación en caso de desastre. El *Software* con métodos de instalación rápidos facilita también la reinstalación en caso de contingencia.

Existe un *Software* que es conocido por la cantidad de agujeros de seguridad que introduce. Se pueden buscar alternativas que proporcionen iguales funcionalidades pero permitiendo una seguridad extra.

2.7.9.2. Consideraciones de una red

Los puntos de entrada en la red son generalmente el correo, las páginas *Web* y la entrada de ficheros desde discos, o de ordenadores ajenos, como portátiles.

Mantener al máximo el número de recursos de red sólo en modo lectura, impide que ordenadores infectados propaguen virus. En el mismo sentido se pueden reducir los permisos de los usuarios al mínimo.

Se pueden centralizar los datos de forma que detectores de virus en modo *batch* puedan trabajar durante el tiempo inactivo de las máquinas.

Controlar y monitorizar el acceso a *Internet* puede detectar, en fases de recuperación, cómo se ha introducido el virus.

2.7.10. Afirmaciones Erróneas de Seguridad

2.7.10.1. Mi sistema no es importante para un *cracker*

Esta afirmación se basa en la idea de que no introducir contraseñas seguras en una empresa no entraña riesgos pues ¿quién va a querer obtener información mía? Sin embargo, dado que los métodos de contagio se realizan por medio de programas automáticos, desde unas máquinas a otras, estos no distinguen buenos de malos, interesantes de no interesantes, etc. Por tanto, abrir sistemas y dejarlos sin claves es facilitar la vida a los virus.

2.7.10.2. Estoy protegido pues no abro archivos que no conozco

Esto es falso, pues existen múltiples formas de contagio, además los programas realizan acciones sin la supervisión del usuario poniendo en riesgo los sistemas.

2.7.10.3. Como tengo antivirus estoy protegido

En general los programas antivirus no son capaces de detectar todas las posibles formas de contagio existentes, ni las nuevas que pudieran aparecer conforme los ordenadores aumenten las capacidades de comunicación, además los antivirus son vulnerables a desbordamientos de búfer que hacen que la seguridad del sistema operativo se vea más afectada aún.

2.7.10.4. Como dispongo de un firewall no me contagio

Esto únicamente proporciona una limitada capacidad de respuesta. Las formas de infectarse en una red son múltiples. Unas provienen directamente de accesos al sistema (de lo que protege un *firewall*) y otras de conexiones que se realizan (de las que no me protege). Emplear usuarios con altos privilegios para realizar conexiones puede entrañar riesgos, además los firewalls de aplicación (los más usados) no brindan protección suficiente contra el *spoofing*.

2.7.10.5. Servidor *Web* actualizado a la fecha

Puede que este protegido contra ataques directamente hacia el núcleo, pero si alguna de las aplicaciones *Web* está desactualizada, un ataque sobre algún *script* de dicha aplicación puede permitir que el atacante abra una *Shell* y por ende ejecutar comandos.

2.7.11. Organismos Oficiales de seguridad Informática

Existen organismos oficiales encargados de asegurar servicios de prevención de riesgos y asistencia a los tratamientos de incidencias, tales como el CERT/CC (*Computer Emergency Response Team Coordination Center*) del SEI (*Software Engineering Institute*) de la *Carnegie Mellon University* el cual es un centro de alerta y reacción frente a los ataques informáticos, destinados a las empresas o administradores, pero generalmente estas informaciones son accesibles a todo el mundo.

2.7.12. Costos elevados en la Seguridad Total

Hoy es imposible hablar de un sistema ciento por ciento seguro, sencillamente porque el costo de la seguridad total es muy alto. "Por eso las empresas, en general, asumen riesgos: deben optar entre perder un negocio o arriesgarse a ser hackeadas. La cuestión es que, en algunas organizaciones puntuales, tener un sistema de seguridad muy acotado les impediría hacer más negocios", "Si un hacker quiere gastar cien mil dólares en equipos para descifrar una encriptación, lo puede hacer porque es imposible de controlarlo. Y en tratar de evitarlo se podrían gastar millones de dólares".

La solución a medias, entonces, sería acotar todo el espectro de seguridad, en lo que hace a plataformas, procedimientos y estrategias. De esta manera, se puede controlar todo un conjunto de vulnerabilidades, aunque no se logre la seguridad total. Y esto significa ni más ni menos que un gran avance con respecto a unos años atrás.

2.8. MODELO OCULTO DE MÁRKOV

Un modelo oculto de Markov o HMM (*Hidden Markov Model*) es un modelo estadístico en el que se asume que el sistema a modelar es un proceso de Markov

de parámetros desconocidos. El objetivo es determinar los parámetros desconocidos (u *ocultos*, de ahí el nombre) de dicha cadena a partir de los parámetros observables. Los parámetros extraídos se pueden emplear para llevar a cabo sucesivos análisis, por ejemplo en aplicaciones de reconocimiento de patrones. Un HMM se puede considerar como la red bayesiana dinámica más simple.

En un modelo de Markov normal, el estado es visible directamente para el observador, por lo que las probabilidades de transición entre estados son los únicos parámetros. En un modelo oculto de Markov, el estado no es visible directamente, sino que sólo lo son las variables influenciadas por el estado. Cada estado tiene una distribución de probabilidad sobre los posibles símbolos de salida. Consecuentemente, la secuencia de símbolos generada por un HMM proporciona cierta información acerca de la secuencia de estados.

Los modelos ocultos de Markov son especialmente aplicados a reconocimiento de formas temporales, como reconocimiento del habla, de escritura manual, de gestos, etiquetado gramatical o en bioinformática.

Una breve explicación del modelo de Markov es un método de previsión muy fiable sería aquel que analizase la evolución de distintos desarrollos teniendo en cuenta las interrelaciones entre dichos desarrollos e introdujese la variable tiempo.

A partir de un estudio del tipo Delphi, se obtienen como conclusiones las probabilidades y las fechas estimadas de ocurrencia de los eventos del cuestionario. Sin embargo, no se consideran las interrelaciones entre los distintos desarrollos.

El modelo de Markov va a caracterizar el desarrollo secuencial tecnológico mediante dos parámetros probabilísticos: la secuencia de los desarrollos y el

tiempo entre desarrollos sucesivos. Estos dos parámetros se pueden representar con los conceptos transición de estados y tiempo de permanencia en el estado.

Se dice que un proceso es de Markov cuando verifica la propiedad de Markov: la evolución del proceso depende del estado actual y del próximo, y no de anteriores o posteriores.

A partir de un Delphi clásico se pueden extraer los parámetros característicos del modelo de Markov. Con estos parámetros se puede hacer un análisis de los procesos de Markov por ordenador, estudiando el proceso secuencial en el tiempo y hallando la distribución de probabilidades en el tiempo de los desarrollos.

Como consecuencia se obtienen un conjunto de cadenas, denominadas cadenas de Markov, que indican posibles caminos para conseguir un desarrollo tecnológico. Usando este tipo de cadenas, se puede realizar una previsión del futuro en la que se analiza la evolución de distintos desarrollos, teniendo en cuenta las interacciones entre desarrollos e introduciendo el variable tiempo.

2.8.1. Historia

Los modelos ocultos de Markov fueron descritos por primera vez en una serie de artículos estadísticos por Leonard E. Baum y otros autores en la segunda mitad de la década de 1960. Una de las primeras aplicaciones de HMM fue reconocimiento del habla, comenzando en la mitad de la década de 1970.

En la segunda mitad de la década de 1980, los HMM comenzaron a ser aplicados al análisis de secuencias biológicas, en particular de DNA²⁴. Desde entonces, se han hecho ubicuos en el campo de la bioinformática.

²⁴ DNA: DeoxyriboNucleic Acid

2.8.2. Arquitectura de un modelo oculto de Markov

El diagrama que se encuentra más abajo muestra la arquitectura general de un HMM. Cada óvalo representa una variable aleatoria que puede tomar determinados valores. La variable aleatoria $x(t)$ es el valor de la variable oculta en el instante de tiempo t . La variable aleatoria $y(t)$ es el valor de la variable observada en el mismo instante de tiempo t . Las flechas indican dependencias condicionales.

Del diagrama queda claro que el valor de la variable oculta $x(t)$ (en el instante t) *solo* depende del valor de la variable oculta $x(t - 1)$ (en el instante $t - 1$). A esto se le llama propiedad de Markov. De forma similar, el valor de la variable observada $y(t)$ solo depende del valor de la variable oculta $x(t)$ (ambas en el instante t).

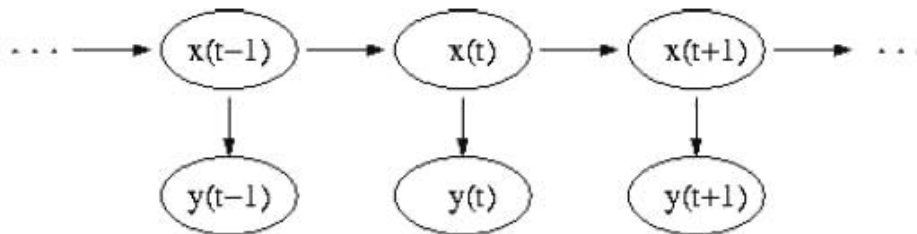


Figura N°2: Arquitectura
Fuente: [Elaboración Propia]

2.8.3. Probabilidad de una secuencia observada

La probabilidad de observar la secuencia

$$Y = y(0), y(1), \dots, y(L - 1)$$

De longitud L está dada por:

$$P(Y) = \sum_X P(Y | X)P(X),$$

Donde la sumatoria se extiende sobre todas las secuencias de nodos ocultos

$$X = x(0), x(1), \dots, x(L - 1).$$

El cálculo por fuerza bruta de $P(Y)$ es impráctico para la mayoría de los problemas reales, dado que el número de secuencias de nodos ocultos será extremadamente alto en tal caso. Sin embargo, el cálculo puede acelerarse notoriamente usando un algoritmo conocido como el procedimiento de avance-retroceso.

2.8.4. Definición formal de un Modelo Oculto de Markov

Una notación habitual de un MOM es la representación como una tupla (Q, V, π, A, B) :

- El conjunto de estados $Q = \{1, 2, \dots, N\}$. El estado inicial se denota como q_t . En el caso de la etiquetación categorial, cada valor de t hace referencia a la posición de la palabra en la oración.
- El conjunto V de posibles valores $\{v_1, v_2, \dots, v_M\}$ observables en cada estado. M es el número de palabras posibles y cada v_k hace referencia a una palabra diferente.
- Las probabilidades iniciales $\pi = \{\pi_i\}$, donde π_i es la probabilidad de que el primer estado sea el estado Q_i .
- El conjunto de probabilidades $A = \{a_{ij}\}$ de transiciones entre estados.
- $a_{ij} = P(q_t = j \mid q_{t-1} = i)$, es decir, a_{ij} es la probabilidad de estar en el estado j en el instante t si en el instante anterior $t - 1$ se estaba en el estado i .

- El conjunto de probabilidades $B = \{b_j(v_k)\}$ de las observaciones.
- $b_j(v_k) = P(o_t = v_k \mid q_t = j)$, es decir, la probabilidad de observar v_k cuando se está en el estado j en el instante t .

La secuencia de observables se denota como un conjunto

$$O = (o_1, o_2, \dots, o_T).$$

2.8.5. Aplicaciones Modelos Ocultos de Markov

Existen tres problemas canónicos asociados con HMM:

- Dados los parámetros del modelo, compútese la probabilidad de una secuencia de salida en particular. Este problema se resuelve con el algoritmo de avance-retroceso.
 - Dados los parámetros del modelo, encuéntrese la secuencia más probable de estados ocultos que puedan haber generado una secuencia de salida dada. Este problema se resuelve con el algoritmo de Viterbi.
 - Dada una secuencia de salida o un conjunto de tales secuencias, encuéntrese el conjunto de estados de transición y probabilidades de salida más probables. En otras palabras, entrénense a los parámetros del HMM dada una secuencia de datos. Este problema se resuelve con el algoritmo de Baum-Welch.
- Aplicaciones de modelos ocultos de Markov
 - Criptoanálisis
 - Reconocimiento del habla, de gestos y de movimientos corporales, reconocimiento óptico de caracteres

- Traducción automática
 - Seguimiento de partituras musicales
 - Bioinformática y Genómica
-
- Predicción de regiones que codifican proteínas dentro de genomas
 - Modelado de familias de secuencias de proteína o ADN relacionado
 - Predicción de elementos de estructura secundaria en secuencias primarias de proteína

2.9. CRIPTOGRAFÍA

Encriptar o criptografía (del griego *krypto*, «oculto», y *graphos*, «escribir», literalmente «escritura oculta») es el arte o ciencia de cifrar y descifrar información mediante técnicas especiales y se emplea frecuentemente para permitir un intercambio de mensajes que sólo puedan ser leídos por personas a las que van dirigidos y que poseen los medios para descifrarlos.

Con más precisión, cuando se habla de esta área de conocimiento como ciencia, se debería hablar de cristología, que a su vez engloba tanto las técnicas de cifrado, es decir, la criptografía propiamente dicha, como sus técnicas complementarias, entre las cuales se incluye el criptoanálisis, que estudia métodos empleados para romper textos cifrados con objeto de recuperar la información original en ausencia de las claves.

2.9.1. Conceptos

En la jerga de la criptografía, la información original que debe protegerse se denomina texto en claro o texto plano. El cifrado es el proceso de convertir el texto plano en un galimatías ilegible, denominado texto cifrado o criptograma. Por lo general, la aplicación concreta del algoritmo de cifrado (también llamado cifra) se basa en la existencia de una clave: información secreta que adapta el algoritmo de

cifrado para cada uso distinto. Cifra es una antigua palabra árabe para designar el número cero; en la Antigüedad, cuando Europa empezaba a cambiar del sistema de numeración romano al arábigo, se desconocía el cero, por lo que este resultaba misterioso, de ahí probablemente que cifrado signifique misterioso.

Las dos técnicas más sencillas de cifrado, en la criptografía clásica, son la sustitución (que supone el cambio de significado de los elementos básicos del mensaje -las letras, los dígitos o los símbolos-) y la trasposición (que supone una reordenación de los mismos); la gran mayoría de las cifras clásicas son combinaciones de estas dos operaciones básicas.

El descifrado es el proceso inverso que recupera el texto plano a partir del criptograma y la clave. El protocolo criptográfico especifica los detalles de cómo se utilizan los algoritmos y las claves (y otras operaciones primitivas) para conseguir el efecto deseado. El conjunto de protocolos, algoritmos de cifrado, procesos de gestión de claves y actuaciones de los usuarios, es lo que constituyen en conjunto un criptosistema, que es con lo que el usuario final trabaja e interactúa.

Existen dos grandes grupos de cifras: los algoritmos que usan una única *clave* tanto en el proceso de cifrado como en el de descifrado, y los que emplean una clave para cifrar mensajes y una clave distinta para descifrarlos. Los primeros se denominan cifras simétricas, de clave simétrica o de clave privada, y son la base de los algoritmos de cifrado clásico. Los segundos se denominan cifras asimétricas, de clave asimétrica o de clave pública y forman el núcleo de las técnicas de cifrado modernas.

En el lenguaje cotidiano, la palabra código se usa de forma indistinta con cifra. En la jerga de la criptografía, sin embargo, el término tiene un uso técnico especializado: los códigos son un método de criptografía clásica que consiste en sustituir unidades textuales más o menos largas o complejas, habitualmente palabras o frases, para ocultar el mensaje; por ejemplo, "cielo azul" podría significar

«atacar al amanecer». Por el contrario, las cifras clásicas normalmente sustituyen o reordenan los elementos básicos del mensaje -letras, dígitos o símbolos-; en el ejemplo anterior, «rcnm arcteeaal aaa» sería un criptograma obtenido por transposición. Cuando se usa una técnica de códigos, la información secreta suele recopilarse en un libro de códigos.

Con frecuencia los procesos de cifrado y descifrado se encuentran en la literatura como encriptado y des encriptado, aunque ambos son neologismos erróneos - anglicismos de los términos ingleses *encrypt* y *decrypt*- todavía sin reconocimiento académico. Hay quien hace distinción entre cifrado/descifrado y encriptado/desencriptado según estén hablando de criptografía simétrica o asimétrica, pero la realidad es que la mayoría de los expertos hispanohablantes prefieren evitar ambos neologismos hasta el punto de que el uso de los mismos llega incluso a discernir a los aficionados y novatos en la materia de aquellos que han adquirido más experiencia y profundidad en la misma.

Ideológicamente cifrar equivale a escribir y descifrar a leer lo escrito.

2.9.2. Historia de la criptografía

La historia de la criptografía es larga y abunda en anécdotas. Ya las primeras civilizaciones desarrollaron técnicas para enviar mensajes durante las campañas militares, de forma que si el mensajero era interceptado la información que portaba no corriera el peligro de caer en manos del enemigo. Posiblemente, el primer criptosistema que se conoce fuera documentado por el historiador griego Polibio: un sistema de sustitución basado en la posición de las letras en una tabla. También los romanos utilizaron sistemas de sustitución, siendo el método actualmente conocido como César, porque supuestamente Julio César lo empleó en sus campañas, uno de los más conocidos en la literatura (según algunos autores, en realidad Julio César no usaba este sistema de sustitución, pero la atribución tiene tanto arraigo que el nombre de este método de sustitución ha quedado para los anales de la historia). Otro de los métodos criptográficos utilizados por los griegos fue la

espartana, un método de trasposición basado en un cilindro que servía como clave en el que se enrollaba el mensaje para poder cifrar y descifrar.

En 1465 el italiano Leon Battista Alberti inventó un nuevo sistema de sustitución poli alfabética que supuso un gran avance de la época. Otro de los criptógrafos más importantes del siglo XVI fue el francés Blaise de Vigenère que escribió un importante tratado sobre "la escritura secreta" y que diseñó una cifra que ha llegado a nuestros días asociada a su nombre. A Selenus se le debe la obra criptográfica "*Cryptomenytices et Cryptographiae*" (Luneburgo, 1624). Durante los siglos XVII, XVIII y XIX, el interés de los monarcas por la criptografía fue notable. Las tropas de Felipe II emplearon durante mucho tiempo una cifra con un alfabeto de más de 500 símbolos que los matemáticos del rey consideraban inexpugnable. Cuando el matemático francés François Viète consiguió criptoanalizar aquel sistema para el rey de Francia, a la sazón Enrique IV, el conocimiento mostrado por el rey francés impulsó una queja de la corte española ante del papa Pío V acusando a Enrique IV de utilizar magia negra para vencer a sus ejércitos. Por su parte, la reina María Estuardo, reina de Escocia, fue ejecutada por su prima Isabel I de Inglaterra al descubrirse un complot de aquella tras un criptoanálisis exitoso por parte de los matemáticos de Isabel.

Durante la Primera Guerra Mundial, los alemanes usaron el cifrado ADFGVX. Este método de cifrado es similar a la del tablero de ajedrez Polibio. Consistía en una matriz de 6 x 6 utilizado para sustituir cualquier letra del alfabeto y los números 0 a 9 con un par de letras que consiste de A, D, F, G, V, o X.



Figura N°3: Máquina Enigma Utilizada por los alemanes durante la II guerra Mundial
Fuente: [W13-2009]

Desde el siglo XIX y hasta la Segunda Guerra Mundial, las figuras más importantes fueron la del holandés Auguste Kerckhoffs y la del prusiano Friedrich Kasiski. Pero es en el siglo XX cuando la historia de la criptografía vuelve a experimentar importantes avances. En especial durante las dos contiendas bélicas que marcaron al siglo: la Gran Guerra y la Segunda Guerra Mundial. A partir del siglo XX, la criptografía usa una nueva herramienta que permitirá conseguir mejores y más seguras cifras: las máquinas de cálculo. La más conocida de las máquinas de cifrado posiblemente sea la máquina alemana Enigma: una máquina de rotores que automatizaba considerablemente los cálculos que era necesario realizar para las operaciones de cifrado y descifrado de mensajes. Para vencer al ingenio alemán, fue necesario el concurso de los mejores matemáticos de la época y un gran esfuerzo computacional. No en vano, los mayores avances tanto en el campo de la criptografía como en el del criptoanálisis no empezaron hasta entonces.

Tras la conclusión de la Segunda Guerra Mundial, la criptografía tiene un desarrollo teórico importante, siendo Claude Shannon y sus investigaciones sobre teoría de la información esenciales hitos en dicho desarrollo. Además, los avances en computación automática suponen tanto una amenaza para los sistemas existentes como una oportunidad para el desarrollo de nuevos sistemas. A mediados de los

años 70, el Departamento de Normas y Estándares norteamericano publica el primer diseño lógico de un cifrado que estaría llamado a ser el principal sistema criptográfico de finales de siglo: el Estándar de Cifrado de Datos o DES. En esas mismas fechas ya se empezaba a gestar lo que sería la, hasta ahora, última revolución de la criptografía teórica y práctica: los sistemas asimétricos. Estos sistemas supusieron un salto cualitativo importante, ya que permitieron introducir la criptografía en otros campos que hoy día son esenciales, como el de la firma digital.

2.9.3. Ramas derivadas

- Criptología
- Criptografía simétrica o convencional
- Criptografía asimétrica o de clave pública
- Criptografía de curva elíptica
- Criptografía híbrida
- Criptografía (música)
- Derecho de las TICs²⁵
- Firma digital
- Esteganografía
- Criptoanálisis
- Infraestructura de clave pública
- Especificaciones PKCS²⁶
- Atbash²⁷
- Test de prioridad

2.9.4. Algoritmos

- Advanced Encryption Standard

²⁵ TICs: Tecnologías de la información y la comunicación

²⁶ PKCS: Estándares de criptografía de clave pública

²⁷ ATBASH: Es un método muy común de codificación del Alfabeto hebreo

- ARC4
- CuaimaCrypt
- DES / TripleDES
- DSA
- ECDSA
- Enigma
- IDEA
- RSA
- TEA / XTEA
- Blowfish

2.9.5. Protocolos

- TLS (*Transport Layer Security TLS/SSL*)
- SSL (*Secure Socket Layer*)
- SET (*Secure Electronic Transaction*)
- OpenPGP (*Pretty Good Privacy o PGP*)
- DSS (*Decision support system*)
- SSH (*Secure SHell, en español: intérprete de órdenes seguro*)

2.9.6. Aplicaciones

- *Software*
 - GNU Privacy Guard, GnuPG o GPG
 - John the Ripper
 - PGP (*Pretty Good Privacy*)
 - WinCuaimaCrypt (Algoritmo criptográfico de llave simétrica)
 - Cifrado de Discos duros y particiones
 - FreeOTFE
 - PointSec
 - Safeboot

- SafeguardDisk
- TrueCrypt
- Dm-crypt
- Voto electrónico
- Pagos electrónicos
 - Transacciones seguras
 - Monedero electrónico

2.10. PUERTOS

2.10.1. Introducción

En informática, un puerto es una forma genérica de denominar a una interfaz a través de la cual los diferentes tipos de datos se pueden enviar y recibir. Dicha interfaz puede ser física, o puede ser a nivel de *Software* (por ejemplo, los puertos que permiten la transmisión de datos entre diferentes ordenadores) (ver más abajo para más detalles), en cuyo caso se usa frecuentemente el término puerto lógico.

2.10.2. Puerto lógico

Se denomina así a una zona, o localización, de la memoria de un ordenador que se asocia con un puerto físico o con un canal de comunicación, y que proporciona un espacio para el almacenamiento temporal de la información que se va a transferir entre la localización de memoria y el canal de comunicación.

En el ámbito de *Internet*, un puerto es el valor que se usa, en el modelo de la capa de transporte, para distinguir entre las múltiples aplicaciones que se pueden conectar al mismo host, o puesto.

Aunque muchos de los puertos se asignan de manera arbitraria, ciertos puertos se asignan, por convenio, a ciertas aplicaciones particulares o servicios de carácter

universal. De hecho, la IANA (*Internet Assigned Numbers Authority*) determina, las asignaciones de todos los puertos comprendidos entre los valores [0 - 1023], (hasta hace poco, la IANA sólo controlaba los valores desde el 0 al 255). Por ejemplo, el servicio de conexión remota telnet, usado en *Internet* se asocia al puerto 23. Por tanto, existe una tabla de puertos asignados en este rango de valores. Los servicios y las aplicaciones que se encuentran en el listado denominado SPA (*Selected Port Assignments*). De manera análoga, los puertos numerados en el intervalo [1024 - 65535] se pueden registrar con el consenso de la IANA, vendedores de *Software* y organizaciones. Por ejemplo, el puerto 1352 se asigna a *Lotus Notes*.

2.10.3. Puerto Físico

Un puerto físico, es aquella interfaz, o conexión entre dispositivos, que permite conectar físicamente distintos tipos de dispositivos como monitores, impresoras, escáneres, discos duros externos, cámaras digitales, memorias *pendrive*. Estas conexiones tienen denominaciones particulares como, por ejemplo, los puertos "serie" y "paralelo" de un ordenador.

Puerto serie (o serial): Un puerto serie es una interfaz de comunicaciones entre ordenadores y periféricos en donde la información es transmitida bit a bit de manera secuencial, es decir, enviando un solo bit a la vez (en contraste con el puerto paralelo que envía varios bits a la vez).

El puerto serie por excelencia es el RS-232²⁸ que utiliza cableado simple desde 3 hilos hasta 25 y que conecta ordenadores o micro controladores a todo tipo de periféricos, desde terminales a impresoras y módems pasando por ratones.

La interfaz entre el RS-232 y el microprocesador generalmente se realiza mediante el integrado 82C50. El RS-232 original tenía un conector tipo D de 25 pines, sin embargo, la mayoría de dichos pines no se utilizaban por lo que IBM (*International*

²⁸ RS-232: *Recommended Standard 232*

Business Machines) incorporó desde su PS/2 (Tipo de conector que es generalmente utilizado para conectar el teclado y el mouse en las PC) un conector más pequeño de solamente 9 pines, que es el que actualmente se utiliza.

En Europa la norma RS-422²⁹, de origen alemán, es también un estándar muy usado en el ámbito industrial.

Uno de los defectos de los puertos serie iniciales era su lentitud en comparación con los puertos paralelos, sin embargo, con el paso del tiempo, han ido apareciendo multitud de puertos serie con una alta velocidad que los hace muy interesantes ya que tienen la ventaja de un menor cableado y solucionan el problema de la velocidad con un mayor apantallamiento. Son más baratos ya que usan la técnica del par trenzado; por ello, el puerto RS-232 e incluso multitud de puertos paralelos están siendo reemplazados por nuevos puertos serie como el USB, el *Firewire* o el Serial ATA (*Advanced Technology Attachment*).

Los puertos serie sirven para comunicar al ordenador con la impresora, el ratón o el módem, sin embargo, el puerto USB sirve para todo tipo de periféricos, desde ratones a discos duros externos, pasando por conexiones *bluetooth*.

Los puertos SATA³⁰: Tienen la misma función que los IDE (*Integrated Drive Electronics*), (a éstos se conecta, la disquetera, el disco duro, lector/grabador de CD y DVD) pero los SATA cuentan con una mayor velocidad de transferencia de datos. Un puerto de red puede ser puerto serie o puerto paralelo.

PCI³¹ : Son ranuras de expansión de la placa madre de un ordenador en las que se pueden conectar tarjetas de sonido, de vídeo, de red, etc. El slot PCI se sigue usando hoy en día y podemos encontrar bastantes componentes (la mayoría) en el

²⁹ RS-422: Protocolo de comunicación de datos en serie que especifica comunicaciones de cuatro cables, dúplex completo, línea diferencial y con varias segregaciones

³⁰ SATA: *Serial ATA o Serial Advanced Technology Attachment*

³¹ PCI: *Interconexión de Componentes Periféricos*

formato PCI. Dentro de los slots PCI está el PCI-Express (Entradas/Salidas de Tercera Generación). Los componentes que suelen estar disponibles en este tipo de slot son:

- Capturadoras de televisión.
- Controladoras RAID (*Redundant Array of Independent Disks*).
- Tarjetas de red, inalámbricas, o no.
- Tarjetas de sonido.

PCI-Express: Es un nuevo desarrollo del bus PCI que usa los conceptos de programación y los estándares de comunicación existentes, pero se basa en un sistema de comunicación serie mucho más rápido que PCI y AGP³². Este sistema es apoyado, principalmente, por Intel, que empezó a desarrollar el estándar con el nombre de proyecto *Arapahoe* después de retirarse del sistema *Infiniband*. Tiene velocidad de transferencia de 16x (8GB/s) y se utiliza en tarjetas gráficas.

Puertos de memoria: A estos puertos se conectan las tarjetas de memoria RAM. Los puertos de memoria son aquellos puertos, o bahías, donde se pueden insertar nuevas tarjetas de memoria, con la finalidad de extender la capacidad de la misma.

Existen bahías que permiten diversas capacidades de almacenamiento que van desde los 256MB (*Megabytes*) hasta 4GB (*Gigabytes*). Conviene recordar que en la memoria RAM es de tipo volátil, es decir, si se apaga repentinamente el ordenador los datos almacenados en la misma se pierden. Dicha memoria está conectada con la CPU a través de buses de muy alta velocidad. De esta manera, los datos ahí almacenados, se intercambian con el procesador a una velocidad unas 1000 veces más rápida que con el disco duro.

³² AGP: *Accelerated Graphics Port*

Puertos inalámbricos: Las conexiones en este tipo de puertos se hacen, sin necesidad de cables, a través de la conexión entre un emisor y un receptor utilizando ondas electromagnéticas. Si la frecuencia de la onda, usada en la conexión, se encuentra en el espectro de infrarrojos se denomina puerto infrarrojo. Si la frecuencia usada en la conexión es la usual en las radio frecuencias entonces sería un puerto *Bluetooth* (redes de área personal inalámbricas).

La ventaja de esta última conexión es que el emisor y el receptor no tienen que estar orientados el uno con respecto al otro para que se establezca la conexión. Esto no ocurre con el puerto de infrarrojos. En este caso los dispositivos tienen que "verse" mutuamente, y no debe interponer ningún objeto entre ambos ya que se interrumpe la conexión.

Puerto USB: Un puerto permite conectar hasta 127 dispositivos y ya es un estándar en los ordenadores de última generación, que incluyen al menos cuatro puertos USB 2.0 en los más modernos, y algún USB 1.1 en los más anticuados

Pero ¿qué otras ventajas ofrece este puerto? Es totalmente *Plug & Play*, es decir, con sólo conectar el dispositivo y "en caliente" (con el ordenador ya encendido), el dispositivo es reconocido, e instalado, de manera inmediata. Sólo es necesario que el Sistema Operativo lleve incluido el correspondiente controlador o driver.

Presenta una alta velocidad de transferencia en comparación con otro tipo de puertos. USB 1.1 alcanza los 12 Mb/s y hasta los 480 Mb/s (60 MB/s) para USB 2.0, mientras un puerto serie o paralelo tiene una velocidad de transferencia inferior a 1 Mb/s. El puerto USB 2.0 es compatible con los dispositivos USB 1.1.

A través del cable USB no sólo se transfieren datos; además es posible alimentar dispositivos externos a través de él. El consumo máximo de este controlador es de 5 voltios.

Los dispositivos se pueden dividir en dispositivos de bajo consumo (hasta 100 mA) y dispositivos de alto consumo (hasta 500 mA). Para dispositivos que necesiten más de 500 mA será necesaria alimentación externa.

Hay que tener en cuenta, además, que si se utiliza un concentrador y éste está alimentado, no será necesario realizar consumo del bus. Una de las limitaciones de este tipo de conexiones es que longitud del cable no debe superar los 5 ms y que éste debe cumplir las especificaciones del *Standard* USB iguales para la 1.1 y la 2.0

2.11. LEYES

2.11.1. Ley 1768 del Código Penal, 2 artículos de Delitos Informático

Ley N° 1768, Ley de 10 de Marzo de 1997, de Gonzalo Sánchez de Lozada, Presidente Constitucional de la República, Código Penal – Reformas.- por cuanto, el Honorable Congreso Nacional, ha sancionado la siguiente Ley. El Honorable Congreso Nacional, decreta: Ley de modificaciones al código penal, en su artículo dos el punto:

57. Incluyese como Capítulo XI, del Título XII, del Libro Segundo del Código Penal, el siguiente: "DELITOS INFORMÁTICOS"

Incluyese como artículo 363 bis, del Código Penal, el siguiente:

(MANIPULACIÓN INFORMÁTICA). - El que con la intención de obtener un beneficio indebido para sí o un tercero, manipule un procesamiento o transferencia de datos informáticos que conduzca a un resultado incorrecto o evite un proceso tal cuyo resultado habría sido correcto, ocasionando de esta manera una transferencia patrimonial en perjuicio de tercero, será sancionado con reclusión de uno a cinco años y con multa de sesenta a doscientos días.

Incluyese como artículo 363 ter del Código Penal, el siguiente:

(ALTERACIÓN, ACCESO Y USO INDEBIDO DE DATOS INFORMÁTICOS).- El que sin estar autorizado se apodere, acceda, utilice, modifique, suprima o inutilice, datos almacenados en una computadora o en cualquier soporte informático, ocasionando perjuicio al titular de la información, será sancionado con prestación de trabajo hasta un año o multa hasta doscientos días.

2.11.2. Reglamento de soporte lógico, D.S. 24582

Decreto Supremo N°24582, Gonzalo Sánchez De Lozada, Presidente constitucional de la República. Considerando:

Que nuestro país cuenta con un ordenamiento jurídico que regula la protección y defensa de los derechos de autor, para cuyo efecto ha sido dictada la Ley 1322 de 13 de abril de 1992, así como el Decreto Supremo Nro. 23907 y otras disposiciones administrativas que rigen la materia en todo el territorio nacional.

Que la Dirección Nacional del Derecho de Autor dependiente de la Secretaría Nacional de Cultura, como organismo directamente vinculado con el ámbito del Derecho de Autor, ha comprobado que existe la imperativa necesidad de definir el régimen de protección del soporte lógico y los bancos de datos así como de regular las relaciones de su explotación en el territorio nacional.

Que, la Ley de Derecho de Autor 1322 de 13 de abril de 1992 en su Art. 6 inciso l) dispone que los programas de ordenador se encuentran protegidos y que requieren de una reglamentación específica.

EN CONCEJO DE MINISTROS, DECRETA:

ARTÍCULO ÚNICO.- Apruébase el Reglamento del soporte lógico o *Software* en sus IX capítulos y 27 artículos dejando claramente establecido que las disposiciones de la Ley de Derecho de Autor y su Decreto Reglamentario son aplicables en su integridad a todas las relaciones jurídicas que se vinculan con los programas de ordenador.

Los señores Ministros de Estado en los despachos de Desarrollo Humano, de Desarrollo Económico, de Hacienda y de Trabajo, quedan encargados de la ejecución y cumplimiento del presente Decreto Supremo.

Es dado en Palacio de Gobierno de la ciudad de La Paz a los veinticinco días del mes de abril de mil novecientos noventa y siete años.

Fdo. Gonzalo Sánchez de Lozada, Antonio Aranibar Quiroga, Víctor Hugo Canelas Zannier, Alfonso Erwin Kreidler Guillaux, José Guillermo Justiniano Sandoval, René Oswaldo Blattmann Bauer, Fernando Candia Castillo, Franklin Anaya Vásquez, Moisés Jarmúsz Levy, Alberto Vargas Covarrubias, Mauricio Antezana Villegas, Alfonso Revollo Thenier, Jaime Villalobos Sanjinés.

Reglamento de *Software*, Capítulo I, Objetivo y definiciones:

Artículo 1.- Objetivo.- De conformidad al inciso "I", artículo 6 de la Ley de Derecho de Autor, de 13 de abril de 1992, el presente reglamento regula los derechos de los autores y titulares de derechos de autor, y define el régimen de protección del soporte lógico y las relaciones de explotación del mismo. El derecho de autor nace con la creación de la obra, de acuerdo a lo previsto en el Art. 2 de la Ley 1322.

De acuerdo al inciso "b", artículo 7, de la mencionada Ley, este reglamento protege también los bancos de datos, considerándolos análogos a las obras derivadas.

Los programas de ordenador y las bases de datos serán protegidos como obras literarias. Constituyéndose en obras intelectuales y formas de expresión creativa del intelecto humano sujetos de protección conforme lo establece en la Decisión 351 del Acuerdo de Cartagena, los ADPIC³³ de la Organización Mundial del Comercio y el Convenio de Berna.

Artículo 2.- Definiciones.- Para los efectos del presente reglamento y la mejor comprensión de los vocablos técnicos en él incluidos, se establecen las siguientes definiciones, las cuales podrán ser actualizadas mediante normas técnicas:

Algoritmo.- Conjunto predeterminado de instrucciones para resolver un problema específico en un número finito de pasos (compárese con Heurística).

Banco de datos.- Conjunto organizado de información accesible por computadora.

Computadora.- (Ordenador).- Dispositivo electrónico que puede almacenar y procesar información.

Copia de respaldo.- ("*Bakup*").- Es la copia del soporte lógico o banco de datos para fines de salvaguarda.

Diagrama de flujo.- Conjunto de símbolos y líneas interconectadas para mostrar un sistema de procesamiento de información o una secuencia de operaciones en programas.

Heurística.- Método de ensayo y error que se vale de reglas empíricas para encontrar la solución para un problema evaluando por etapas los progresos hechos a lo largo de su curso. (Compárese con Algoritmo).

Lenguaje de programación.- Conjunto de instrucciones con semántica y sintaxis, con los cuales se puede desarrollar un programa fuente.

³³ ADPIC: Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio

Licencia de uso.- Documento mediante el cual se otorga autorización de uso no-exclusivo y no-transferible del soporte lógico o *Software*, de acuerdo a los términos y condiciones mencionadas en el presente reglamento. Este convenio de licencia le permite a un solo usuario instalar el *Software* en una sola computadora y un solo lugar y una sola vez, excepto acuerdo tácito en el mismo que amplíe dichas condiciones.

Memoria.- Dispositivo capaz de recibir datos, retenerlos y suministrarlos a requerimiento del usuario.

Programa para computadora.- Conjunto de instrucciones para ser usadas, directa o indirectamente, en una computadora a fin de obtener un resultado determinado.

Programa fuente.- Conjunto de instrucciones para ser usadas, directa o indirectamente en una computadora a fin de obtener un resultado determinado, en el lenguaje comprensible en el ser humano.

Soporte lógico.- (*Software*).- El soporte lógico es un conjunto de uno o varios programas para computadora, puede incluir información de apoyo, documentación y material auxiliar, cualquiera sea su forma de expresión y fijación.

Soporte informático.- Todo dispositivo o medio físico (memoria, disquetes, discos duros, cintas, etc.) o medio magnético, óptico, químico o papel y otros, empleado para propósitos de comunicación entre humanos y máquinas y fines de almacenamiento.

Soporte físico.- *Hardware* Comprende la totalidad de dispositivos mecánicos, magnéticos, eléctricos y electrónicos en una instalación de procesamiento de datos.

Capítulo II, De la protección

Artículo 3.- Obras protegidas.- De conformidad al art. 4 de la Ley No 1322, el presente reglamento protege el derecho de autor sobre el soporte lógico y los bancos de datos, que con características de individualidad y originalidad surgen y se exteriorizan en una forma de expresión susceptible de ser reproducida e incorporada en un soporte informático, sin extenderse a las ideas, al procedimiento, al lenguaje de programación usados o incluidos en dicha obra.

Los derechos reconocidos al autor son independientes de la propiedad del objeto corporal que contiene la obra.

Artículo 4.- Derechos Morales.- Los Derechos Morales de los autores de soporte lógico están protegidos por el art. 14 de la Ley de Derecho de Autor. Por la vía de excepción y sin vulnerar los derechos morales, este reglamento amparado por el art. 6, inc. i), de la Ley de Derecho de Autor permite modificaciones y mejora el soporte lógico y el banco de datos.

Artículo 5.- Derechos Patrimoniales.- De conformidad con lo dispuesto en los artículos 15 y 17 de la Ley 1322, solamente los titulares de los derechos patrimoniales en soportes lógicos pueden autorizar o prohibir toda forma de explotación de los mismos, en particular su comercialización, arrendamiento, su difusión, reproducción, adaptación, modificación, mejoras, traducción, transformación y la importación.

Artículo 6.- Transferencia del Soporte Informático.- La transferencia del soporte informático que contiene el soporte lógico y el banco de datos otorgan al adquirente el derecho de uso y explotación únicamente en el marco de la licencia de uso.

Artículo 7.- Comunidad Ganancialicia.- En la comunidad ganancialicia el cónyuge, autor de obras de soporte lógico y/o banco de datos, conservará su derecho moral y patrimonial conforme a lo establecido por el art. 107 del Código de Familia.

Artículo 8.- Obras derivadas.- El presente reglamento protege también el soporte lógico y el banco de datos derivados, que resulta de la adaptación o transformación de un soporte lógico, siempre que constituya una creación autónoma y posea originalidad, sin perjuicio de los derechos de autor sobre dicha obra.

Artículo 9.- Secreto-Autoral- Las especificaciones del soporte lógico, los algoritmos, los programas fuente, el diseño del producto, los diagramas de flujo, heurísticas y demás medios de creación del soporte lógico, constituyen secreto autoral y el autor y/o titular no está obligado a revelar tales elementos.

Capítulo III: De los convenios y contratos

Artículo 10.- Licencia de Uso.- Contrato de adhesión mediante el cual el Titular de los derechos de autor, otorga una licencia de Uso.

Artículo 11.- Convenios o Contratos.- La transferencia de los Derechos Patrimoniales se efectuarán mediante convenios o contratos en el marco de lo establecido por el artículo 29 de la Ley de Derecho de Autor y deberán ser registrados de acuerdo a lo establecido por el artículo 26 del Decreto Supremo Reglamentario 23907 de 7 de diciembre de 1994.

Artículo 12.- Obras por Encargo.- El soporte lógico y el banco de datos que se cree bajo un contrato laboral o de prestación de servicios y/o el que fuera desarrollado por empleados o funcionarios públicos en cumplimiento de las obligaciones inherentes a sus cargos, tendrán como titular a la persona natural o jurídica por cuya cuenta y riesgo se realizan, salvo que exista un convenio o contrato que indique lo contrario de conformidad al artículo 29 de la Ley 1322.

Capítulo IV: Protección al derecho de autor

Artículo 13.- Aplicabilidad penal.- En los casos de violación al Derecho de Autor, se aplicarán las normas establecidas en el título XIV, Capítulo I, de la Ley No. 1322; y el capítulo X y XI de la Ley de Modificaciones del Código Penal en sus artículos 362 y 363.

Artículo 14.- Ejemplares ilícitos.- Se registrarán por lo previsto en la Ley No. 1322 y el Decreto Supremo No. 23907.

Artículo 15.- Copia de respaldo.- El usuario que haya adquirido legalmente el derecho de utilización de un soporte lógico o de un banco de datos podrá, excepcionalmente y por sus propios medios, producir su copia de respaldo. El destino de esta copia no es el uso o explotación, si no garantizar la continuación del uso en caso de daño del soporte informático que contiene originalmente con el soporte lógico.

Artículo 16.- Carga de programas.- La carga de programa como paso necesario de su ejecución, por quien se halla legítimamente autorizado mediante una licencia de uso del soporte lógico y/o banco de datos, no constituye acto ilícito.

Capítulo V: de las medidas precautorias, jurisdiccionales y de los medios probatorios

Artículo 17.- Medidas precautorias.- Con carácter provisional y accesorio, se podrán solicitar todas las medidas precautorias que la ley permite.

Artículo 18.- Medidas jurisdiccionales.- De conformidad a lo establecido en el Código de Procedimiento Penal, en sus artículos 190 al 193, la autoridad competente a solicitud de parte interesada podrá disponer: anotación preventiva, requisa, allanamiento, secuestro, precintado, arraigo y toda medida que la ley lo permita.

Artículo 19.- Medidas probatorias.- Son válidas todas las medidas probatorias reconocidas por el ordenamiento jurídico vigente.

Capítulo VI: De la sociedad autoral

Artículo 20.- Reconocimiento.- De conformidad con el artículo 64 de la Ley de Derecho de Autor y el artículo 27 numeral "2" inciso f) del Decreto Supremo No. 23907, se podrá constituir la sociedad de derecho de autor de creadores de programas de ordenador o computadora (soporte lógico o *Software*), previo reconocimiento de la Dirección Nacional de Derecho de Autor.

Artículo 21.- Atribuciones.- Son atribuciones de la sociedad autoral además de las establecidas por el artículo 27 del Decreto Supremo No. 23907, las siguientes:

a.- Asesorar a los titulares de una obra sobre las condiciones a las que deberán ajustarse los contratos con los usuarios, siempre que no contravenga a lo dispuesto por la Ley No. 1322, el Decreto Supremo No. 23907 y el presente reglamento.

b.- Llevar a cabo toda otra acción necesaria por ante las instancias correspondientes para lograr la correcta aplicación y cumplimiento de la ley y su reglamentación.

Capítulo VII: Del registro del soporte lógico, y del banco de datos

Artículo 22.- Registro.- El registro del Soporte Lógico se efectuará en la Dirección Nacional de Derecho de Autor de la Secretaría Nacional de Cultura, dentro del marco de los reglamentos y requisitos vigentes. La Dirección Nacional de Derecho de Autor es responsable de la custodia y de la guarda de la información que se le confía, por lo que bajo ningún concepto podrá develar a terceros sin previa orden judicial debidamente justificado el derecho o interés.

La Resolución Administrativa de registro no es constitutiva de derechos y se otorgará presumiendo la buena fe del solicitante reservando el derecho de terceros.

Artículo 23.- Objetivos.- Son objetivos del registro del soporte lógico, los siguientes:

- a.- Brindar una mayor seguridad del derecho registrado.
- b.- Dar publicidad al derecho de los titulares y a los actos y contratos que transfieran o cambien ese dominio amparado por ley, dando a conocer en beneficio del autor, la existencia de creaciones protegidas por el derecho de autor como medio para demostrar la titularidad sobre la misma.
- c.- Permitir a los usuarios que tengan interés en su explotación, informarse sobre sus condiciones jurídicas para una posible contratación y adicionalmente enterarse de la existencia de los titulares y demás causahabientes, que en virtud de un acto intervivos o por causa de muerte hayan adquirido legítimamente los derechos patrimoniales y en consecuencia tengan viabilidad en su disposición.

Artículo 24.- Materia a registrarse.- En la Dirección Nacional de Derecho de Autor, deberán inscribirse además de los establecidos por el artículo 26 del Decreto Supremo No. 23907, y el presente reglamento, las siguientes:

- a.- El directorio que representa a la sociedad de autores del *Software* adjuntando el acta de elección en asamblea general.
- b.- Personalidad jurídica.
- c.- Las demás que establezcan los reglamentos de la Dirección Nacional de Derecho de Autor.

Capítulo VIII: Del procedimiento administrativo, De conciliación y arbitraje

Artículo 25.- Reglamentación vigente.- Dentro la normativa jurídica como medio alternativo de solución de controversias y en estricta aplicación de la Ley 1770 de

10 de marzo de 1997, Ley de Arbitraje y Conciliación, el presente reglamento deberá sujetarse a la misma.

Artículo 26.- Conciliación y arbitraje.- El procedimiento de Arbitraje se sujetará a lo establecido en los artículos 38 al 84, y en lo que concierne a la Conciliación, a los artículos 91 y 92 de la mencionada Ley de Arbitraje.

Capítulo IX: Disposiciones transitorias

Artículo 27.- Que, de acuerdo a la aplicación de las nuevas tecnologías que surgieran en el futuro, así como las necesidades que ellas determinen, el presente reglamento podrá ser modificado conforme lo determine una disposición expresa.



CAPÍTULO 3

ELABORACIÓN DEL MODELO DE DETECCIÓN DE IRRUPCIONES

A la medida que crece y se diversifica el uso de Infraestructuras Tecnológicas, se incrementan también los riesgos de que los equipos de cómputo, dispositivos electrónicos y sistemas informáticos, conectados o no a *Internet*, sean vulnerables a ataques o incidentes que ponen en peligro la integridad, disponibilidad y autenticidad de los datos que en ellos se procesa, almacena o transfiere. Y más allá de los datos, el daño a dichas infraestructuras es latente.

En este Capítulo se plantea un modelo de análisis forense en las Infraestructuras Tecnológicas específicamente para los sistemas de información que cubra los pasos necesarios desde el aseguramiento de la escena del delito hasta la presentación de evidencias.

El modelo muestra la base para el desarrollo del Código de Prácticas hacia una posterior implementación en *Software* comercial, se crea con la idea de que sea lo más abierto posible. Dado que es necesario ya que el mundo de la tecnología informática avanza rápidamente.

Toda evidencia digital es y debe ser convincente ante un Tribunal de Justicia o en donde haya alguna disputa. Para asegurarla, es importante la homogenización del protocolo de admisibilidad de la prueba, además de un continua aproximación de lo que podría tratarse de una prueba. Esta labor se hace muy difícil en circunstancia en donde no exista un modelo de cómo basarse, menos aún, cuando no exista un marco de trabajo.

3.1. DEFINICIÓN DE FASES DEL MODELO

Se define los pasos que ayudan a detectar las irrupciones en informática forense para los sistemas de información, cada paso muestra el comportamiento de sus

variables y las acciones que debe realizar enlazados en secuencia. Se ve a continuación los pasos del mismo:



Figura N°4: Fases de la detección de Irrupciones
Fuente: [Elaboración Propia]

3.1.1. Validación de usuarios

Lo primero en un proceso de investigación para la detección de irrupciones en informática forense para los sistemas de información, al igual que en cualquier otro proceso criminal, es asegurar (restringir el acceso a la zona del delito para no modificar evidencias) la escena del delito informático.

Idealmente, los usuarios deben ser integrados por un experto en Informática y/o informática forense. El proceso de aseguramiento se debe hacer mediante el código de Usuario, aunque la exactitud debe primar sobre la rapidez en todo el proceso, este paso debe ser realizado por una persona “competente” de la organización implicada que pueda explicar los pasos que ha Realizado y la implicación de sus acciones.

Normalmente los administradores de los sistemas informáticos serán los primeros en tener contacto con la escena del delito y junto a equipo de respuesta de incidentes realizarán los primeros pasos para “congelar” la escena del delito informático.

El rol fundamental de las primeras personas en responder al delito es no hacer nada que pueda producir daño. A menos que esté específicamente entrenado en respuesta a incidentes. Es muy fácil que un malicioso criminal informático inserte un troyano o código hostil que destruya evidencias automáticamente.

Es muy importante que una persona sea asignada con autoridad suficiente para tomar decisiones finales que aseguren la escena del delito, conducir las búsquedas de evidencias y preservar las mismas. Este rol normalmente debe ser asumido por el jefe del equipo forense.

El primero paso del modelo se detalla en el siguiente Diagrama:



Figura N°5: Validación de Usuario
Fuente: [Elaboración Propia]



Figura N°6: Planteamiento del problema a resolver
Fuente: [Elaboración Propia]

A1: Identificar la escena del delito informático. Para ello se debe establecer un perímetro o registro de Base de datos para saber quienes hicieron los últimos ingresos al sistema de información.

A2: Planteamiento del Problema a resolver, o realizar una lista con los sistemas involucrados en el delito.

A2.1: Restringir el acceso a los últimos registros del sistema, acceso tanto de personas (usuario) como acceso de otros equipos informáticos hacia el sistema de información.

A2.2: Esquematizar la escena del delito informática forense.

A2.3: Registrar los ingresos mediante las conexiones de red.

A2.4: Comprobar y registrar si existieran las conexiones inalámbricas que puedan permitir la activación de conexiones remotas.

A2.5: Si hay impresoras imprimiendo, dejar que terminen de imprimir.

3.1.2. Identificación de Evidencias

Es el proceso de conocer los datos, dónde están localizados y cómo están almacenados. Al ser un universo tan heterogéneo el de los sistemas de información donde se pueden encontrar evidencias digitales, se hace necesaria una clasificación para poder organizar las mismas.

Se debe realizar una primera distinción entre evidencias volátiles (evidencias que desaparecen pronto debido a falta de alimentación eléctrica, corte de conexiones telemáticas, etc.) y no volátiles (evidencias que perduran aun a falta de alimentación eléctrica, etc.).

El obtener las evidencias volátiles lo más rápidamente posible es fundamental. La obtención de evidencias volátiles se puede dar en los siguientes lugares:

Volátiles: Registros y *cache* del procesador, Tablas de rutas, *Cache* ARP³⁴, Tabla de procesos, Estadísticas del *kernel* y módulos, Memoria RAM³⁵, Ficheros temporales del sistema, Estado de la red, Tiempos de los ficheros MAC³⁶: modificación, acceso y creación)

³⁴ Cache ARP: Tabla que almacena las asignaciones entre la capa de enlace de datos y direcciones de red de la Capa de direcciones.

³⁵ RAM: Random Access Memory traducido a Memoria de Acceso Aleatorio)

³⁶ MAC: Modificación Acceso y Creación

Hasta aquí se ha obtenido evidencias que se podrían perder, incluso sin reiniciar el equipo que afectarían a un sistema de información.

Otras evidencias volátiles que seguro se pierden al reiniciar el funcionamiento del equipo y que se deben guardar son: Sistemas de ficheros³⁷ montados, Sistemas de ficheros virtuales.

Toda evidencia volátil³⁸ conseguida debe ser grabada como fichero a un dispositivo de almacenamiento y fuera del dispositivo donde están las evidencias, preservando su integridad de la fuente. A partir de este momento, las evidencias volátiles tendrán el mismo tratamiento que las evidencias no volátiles³⁹, por lo cual seguirá los mismos pasos del modelo. Si por cualquier circunstancia no se pueden grabar a dispositivos las evidencias volátiles, hay que ver la factibilidad de realizar un estudio en línea de las mismas, con la pérdida de evidencias o pérdida de integridad de las mismas que ello puede suponer. También debemos clasificar las evidencias en la detección de irrupciones en los sistemas de información en función de acuerdo a lo siguiente:

³⁷ Sistema de ficheros: Conjunto algoritmos y estructuras auxiliares que nos van a permitir de manera sencilla y transparente acceder nuestros datos en dispositivos de almacenamiento.

³⁸ Evidencia Volátil: Evidencia cuya información se pierde al interrumpirse

³⁹ Evidencia no Volátil: evidencia donde es guardado aun después de haber existido una interrupción.



Figura N°7: Identificación de Evidencias
Fuente: [Elaboración Propia]

B1: Identificación en Memorias de almacenamiento: Discos duros, disquetes, CDs⁴⁰, DVDs⁴¹. Normalmente en estos dispositivos es donde se puede enviar información referente a las irrupciones en un sistema informático. La evidencia digital estará contenida en los sistemas de ficheros de cada uno de estos dispositivos.

B2: Redes: Por donde se obtiene evidencias: Tarjetas de red de los ordenadores y protocolos de Red

B3: Redes Inalámbricas: donde obtiene evidencias de las tarjetas inalámbricas, Puntos de accesos. Dispositivos móviles. Sistema de Navegación por satélite.

En función de las prioridades del cliente (recuperación de cierta información, saber cómo ocurrió el delito, obtener la pruebas para llevar a juicio al delincuente) el

⁴⁰ CDs: Discos compacto, memoria de sólo lectura

⁴¹ DVDs: Discos Versátil Digital de Alta Densidad, con formato de almacenamiento óptico

experto debe buscar unas evidencias u otras. Dicho experto no debe gastar horas innecesarias recogiendo información que no sea relevante.

Casi siempre las evidencias estarán localizadas en el sistema de fichero del dispositivo o equipo comprometido, por lo cual el experto forense debe realizar una copia a nivel de bits de dicho sistema de ficheros. Un último punto sería recordar que la forma de localizar las evidencias no vaya en contra de ninguna ley del país. Hay que conocer la normativa legal sobre la interceptación de datos de terceros en medios digitales.

3.1.3. Preservar las Evidencias

Esta es la fase más importante y crítica del modelo de detección de irrupciones, puesto que una vez que se halla comprobado el delito informático la empresa o institución dañada normalmente deseará llevar a un proceso judicial al atacante. Para ello es necesario poseer evidencias digitales preservadas de tal forma que no haya duda alguna de su verosimilitud y siempre de acuerdo a las leyes vigentes.

Este proceso de preservación se debe realizar tan pronto como sea posible. Siempre que sea posible hay que evitar los cambios en las evidencias y si no se logra registrarlos, documentarlos y justificarlos, siempre que sea posible con testigos de registro de evidencias que puedan corroborar las acciones. Recordemos que las primeras evidencias que hay que obtener son las volátiles, que al guardarlas en ficheros o una base de datos se convertirán en evidencias no volátiles.

Entonces tomamos las tareas siguientes para las evidencias digitales:

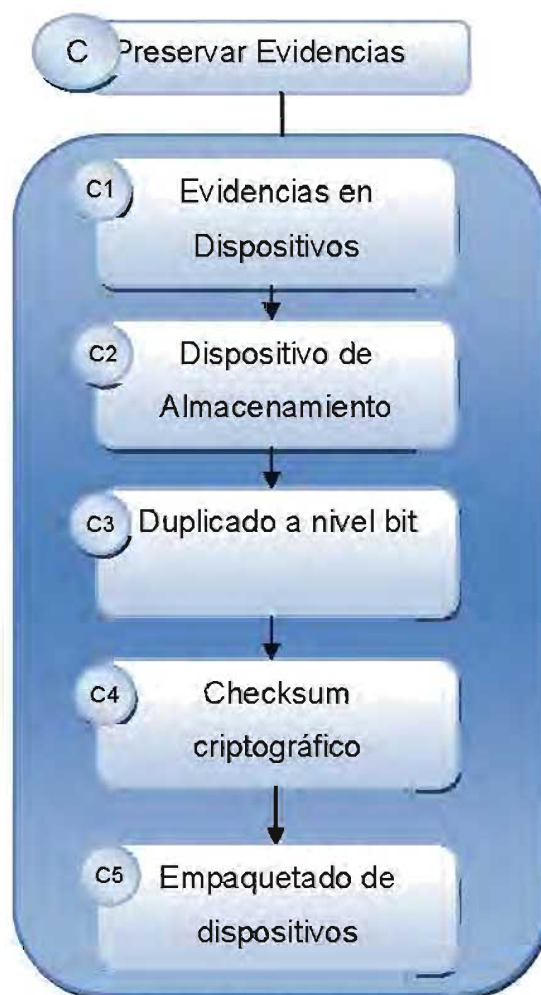


Figura N°8: Preservar Evidencias
Fuente: [Elaboración Propia]

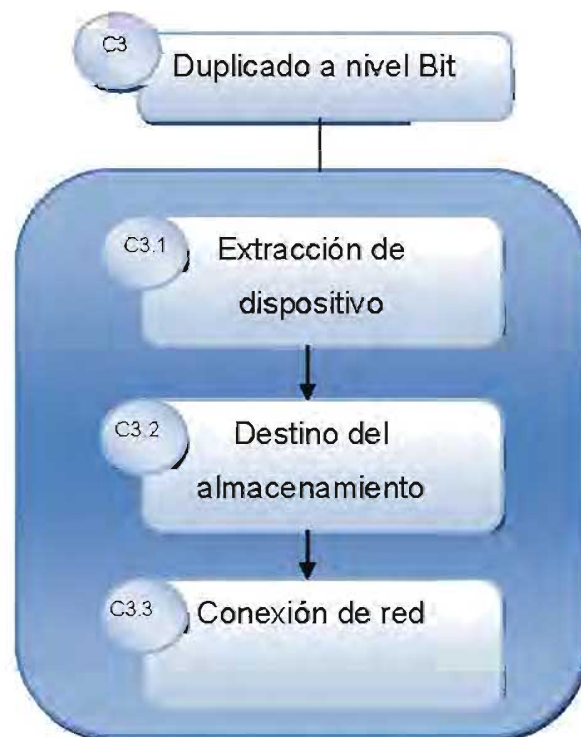


Figura N°9: Duplicado a nivel bit
Fuente: [Elaboración Propia]



Figura N°10: Empaquetado de dispositivos
Fuente: [Elaboración Propia]

C1. Si el dispositivo del cual tenemos que hacer copia de su sistema de almacenamiento está encendido, extraerlo siempre que sea posible y ponerlo en una estación de trabajo para la adquisición de datos.

C2. Toda evidencia digital guardada en dispositivos de almacenamiento, y por tanto almacenado en un sistema de ficheros, debe ser copiado mediante procedimientos del modelo de detección de irrupciones en informática forense que no alteren la evidencia y que sean admisibles en un tribunal de justicia. Para ello realizar una imagen a nivel de bit del sistema de almacenamiento del dispositivo. Una imagen a nivel de bits es una copia que registra cada bit que fue grabado en el dispositivo de almacenamiento original, incluyendo ficheros ocultos, ficheros temporales, ficheros corruptos, ficheros fragmentados y ficheros borrados que todavía no han sido sobrescritos.

C3. Formas para crear duplicados a nivel de bit de los discos de almacenamiento de información.

C3.1 Extraer del dispositivo origen y copiar.

C3.2 Usar un dispositivo destino para el almacenamiento de la información.

C3.3 Usar una conexión de red, conexión *Ethernet*, cable cruzado, USB, para transferir el contenido del disco al otro dispositivo de almacenamiento.

Retención de tiempos y fechas. El tiempo y fecha de creación o modificación de un fichero puede ser un importante asunto en un delito. Si el usuario puede tener el sistema sin configurar apropiadamente el tiempo o deliberadamente cambiar las propiedades de fecha y hora, los ficheros puede que no sean correspondientes con la fecha real. Esto puede ser un problema si, por ejemplo, el sistema de registro muestra que un fichero fue creado en una fecha concreta y el sospechoso es capaz de probar que esa fecha no usó el ordenador. Por ello se debe anotar hora y fecha del sistema antes de apagarlo, documentando el hecho. Además puede ser prudente fotografiar la pantalla mostrando el acceso a ficheros o tiempos de modificación antes de abrir dichos ficheros. También tener en cuenta el desfase

horario que pueda haber entre el dispositivo que contiene la evidencia y el horario real, documentado este desfase.

Siempre que sea posible trabajar con zonas de tiempo GMT. El delito puede involucrar varias zonas de tiempo y usando GMT puede ser un punto de referencia que haga el análisis de las evidencias más sencillo.

C4. Generar los procesos de *checksum*⁴² criptográfico de la copia y del original. Mediante el método de *checksum* criptográfico, proceso de generación de la integridad de un fichero, conjunto de ficheros o de toda la información contenida en un dispositivo de almacenamiento, se garantiza que la evidencia no será alterada en ni un solo *bit*. El proceso es sencillo; generar el *checksum* significa generar un hash, valor único para un determinado conjunto de *bytes*, de la evidencia. Esto es posible dado que los algoritmos criptográficos de hash son cuidadosamente seleccionados para ser funciones de un solo sentido: dado un determinado *checksum* criptográfico para un mensaje, es virtualmente imposible adivinar qué mensaje produjo ese *checksum*. Dicho de otra manera, no es posible hallar mediante cálculos dos mensajes que generen el mismo *checksum* criptográfico. Gracias a determinado *Software* especializado y algoritmos de verificación de *checksum* se comprueba que si la evidencia no se ha alterado produce un hash idéntico al original. También podemos usar firma digital para realizar el proceso de autenticación de la copia y del original, puesto que debido a sus características (única, no falsificable, fácil de autenticar, barata y fácil de generar) es ideal para este proceso.

C5. Empaquetar los dispositivos que contiene las evidencias. Los detalles mínimos que deben ser registrados y directos e inequívocamente atribuidos a cada paquete son:

⁴² Checksum: Es la suma de verificación de tramas, al enviar una trama de datos, se calcula un número sumando los bits.

C5.1 Identificador único

- Breve descripción del material
- Localización desde donde y por quien.
- Día y hora de registro.

C6. Si el paquete debe ser enviado mediante correo o vía red, hay que asegurarse de usar un método que permita el seguimiento del mismo. En este punto la evidencia digital está preservada.

3.1.4. Análisis de las Evidencias

El concepto de evidencia digital se forma (normalmente) por el contenido de los ficheros (datos) y la información sobre los ficheros (metadatos). Basándose en estas evidencias el investigador debe intentar contestar a las siguientes preguntas en la fase de análisis:

¿Quién?

Reunir la información sobre el/los individuo/s involucrados en el compromiso.

¿Qué?

Determinar la naturaleza exacta de los eventos ocurridos.

¿Cuándo?

Reconstruir la secuencia temporal de los hechos.

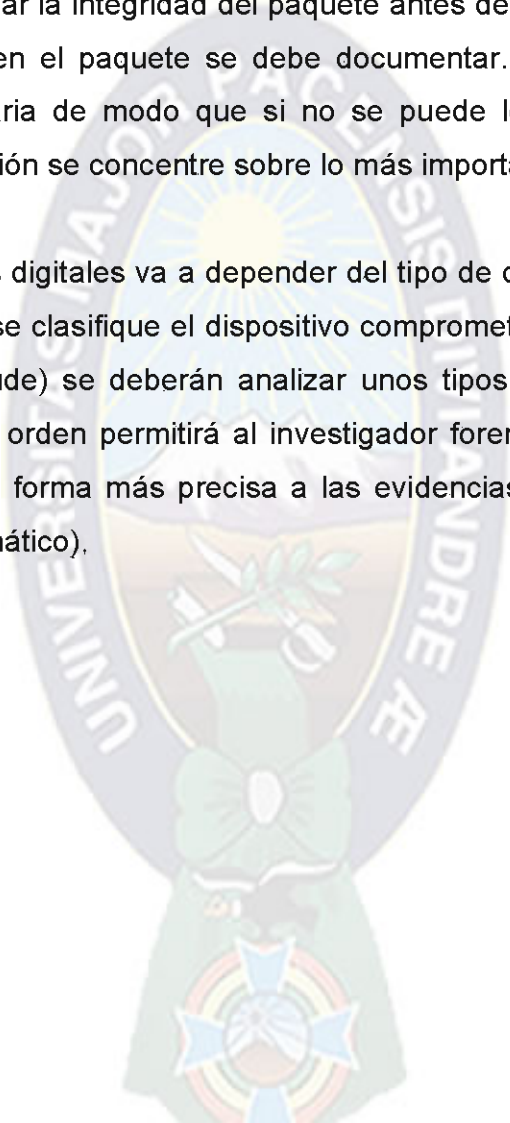
¿Cómo?

Descubrir que herramientas o *exploits* se han usado para cometer el delito.

La evidencia almacenada debe ser analizada para extraer la información relevante y recrear la cadena de eventos sucedidos. El análisis requiere un conocimiento profundo de lo que se está buscando y como obtenerlo.

Cualquier elemento enviado para su análisis forense debería ser en primer lugar revisado para comprobar la integridad del paquete antes de empezar dicho análisis. Cualquier deficiencia en el paquete se debe documentar. Es importante ver que información es prioritaria de modo que si no se puede lograr una recuperación completa, la recuperación se concentre sobre lo más importante.

Analizar las evidencias digitales va a depender del tipo de datos a analizar, del tipo de sistema en el cual se clasifique el dispositivo comprometido. Además en función del tipo de delito (fraude) se deberán analizar unos tipos de evidencias y en un determinado orden (el orden permitirá al investigador forense informático llegar lo antes posibles y de la forma más precisa a las evidencias digitales para llegar a resolver el delito informático).



Existen cuatro categorías de datos:

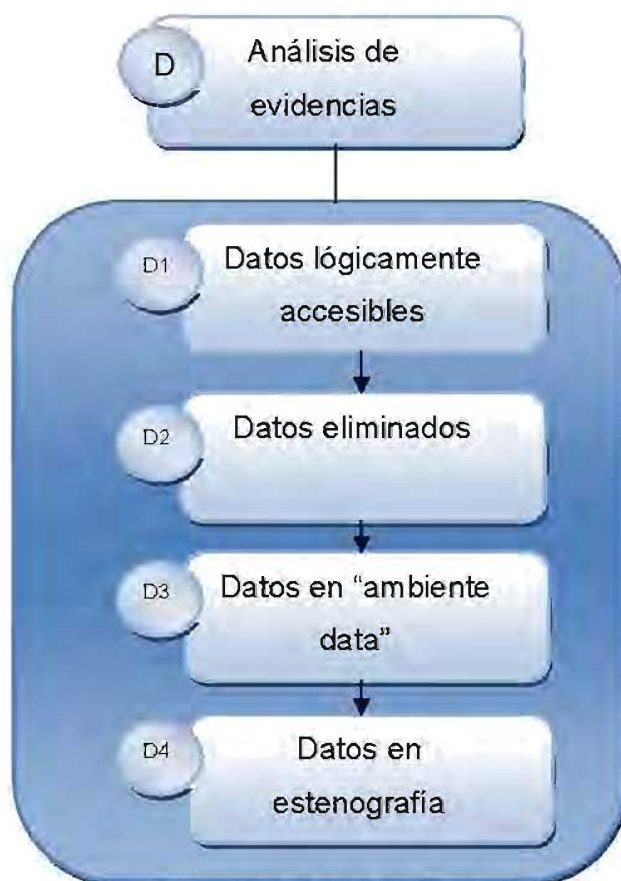


Figura N°11: Análisis de Evidencias
Fuente: [Elaboración Propia]

D1. Datos lógicamente accesibles: Son los datos más comunes. Las dificultades que podemos encontrar en estos Datos son:

Que haya una gran cantidad de información a analizar (los actuales dispositivos de almacenamiento pueden contener una cantidad ingente de ficheros).

- Que estén cifrados. Si están cifrados con programas como por ejemplo *Office* es fácil romper la clave; en otros casos es virtualmente imposible romperlo.

- Que estén corruptos o que tengan trampa, se debe usar buscadores de virus para encontrar una clave maliciosa metida en archivos de evidencias, antes de que puedan crear estragos. Aunque la mayoría de los virus residen en programas ejecutables que raramente serán ejecutados en el transcurso de una investigación forense, los virus aprovechan las vulnerabilidades de algunos Sistemas Operativos y aplicaciones y pueden ser accionados simplemente al ver los documentos en los cuales residen.

D2. Datos que han sido eliminados (si no han sido sobrescritos se pueden volver a recuperar).

D3. Datos en los espacio no asignado, ficheros de *swap/page file*, espacio entre sectores, espacio entre particiones, datos treams alternativos. Este tipo de datos necesita *Software* especial para poder ser recuperados.

D4. Datos en estenografía (proceso por el cual se puede ocultar datos dentro ficheros). Permite detectar la presencia de datos que están ocultos dentro de ficheros usando técnicas de estenografía. Detectar la presencia de estenografía es más fácil que la extracción de los datos ocultos en sí mismo.

A continuación se observa una clasificación de delitos informáticos a estudiar para la fase:

- a) Fraudes cometidos mediante manipulación de ordenadores.
- b) Manipulación de programas.
- c) Manipulación de datos de salida.
- d) Fraude efectuado por manipulación informática o por medio de dispositivos informáticos.
- e) Falsificaciones informáticas.
- f) Sabotaje informático.

- g) Virus, gusanos y bombas lógicas.
- h) Acceso no autorizado al Sistemas o Servicios de Información.
- i) Reproducción no autorizada de programas informáticos de protección legal.
- j) Amenazas mediante correo electrónico.

3.1.5. Presentación é Informes

Basándose en las fases anteriores, en toda la documentación disponible del caso y basándose también en la cadena de custodia, la presentación y/o sustentación del informe pericial es la fase de comunicar el significado de la evidencia digital, los hechos, sus conclusiones y justificar el procedimiento empleado.

El propósito de la presentación de los informes es proporcionar al lector toda la información relevante de las evidencias de forma clara, concisa, estructurada y sin ambigüedad para hacer la tarea de asimilación de la información tan fácil como sea posible.



Figura N°12: Presentación de Informes
Fuente: [Elaboración Propia]

E1. La forma de presentación es muy importante y debe ser entendible por personas no conocedoras del tema en discusión. Es decisivo que el modelo presente las evidencias en un formato sencillo de entender, acompañado de explicaciones que eviten la jerga y la terminología técnica.

3.2. DIAGRAMA GENERAL

3.2.1. Cadena de comportamiento Prioritario



Figura N°13: Cadena de comportamiento
Fuente: [Elaboración Propia]

3.2.2. Comportamiento en la Red

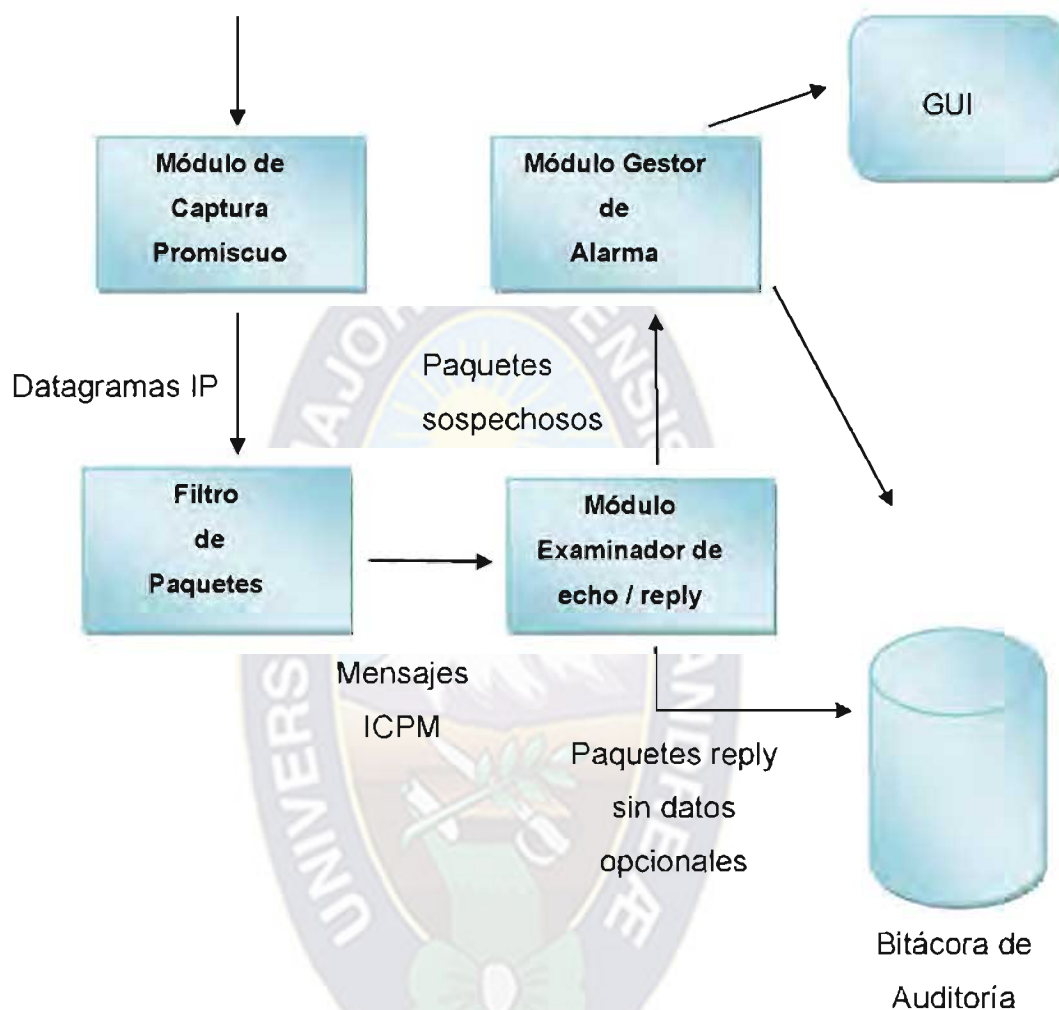


Figura N°14: Comportamiento de la Red
Fuente: [Elaboración Propia]

3.3. MODELADO MEDIANTE MARKOV

A continuación se aplican los Modelos Ocultos de Markov que permitirá detectar las irrupciones en los puertos de la red y guardarlas como evidencia en la detección de irrupciones en la informática forense.

3.3.1. Modelos Ocultos de Markov

N : Número de estados del modelo

Estados, $s = \{s_1, s_2, \dots, s_N\}$ estado en tiempo $t, q_t \in s$

M : Número de símbolos de observación

Símbolos de observación, $v = \{v_1, v_2, \dots, v_M\}$

Observación en tiempo $t, a_t \in v$

$A = \{a_{ij}\}$: Distribución de la probabilidad de la transición del estado

$$a_{ij} = P(q_{t+1} = s_j | q_t = s_i), 1 \leq t, j \leq N$$

$B = \{b_j(k)\}$: Distribución de la probabilidad del símbolo de observación del estado j .

$$b_j(k) = P(v_k a_t t | q_t = s_j), 1 \leq j \leq N, 1 \leq k \leq M$$

$\pi = \{\pi_i\}$: Distribución del estado inicial

$$\pi_i = P(q_1 = s_i), 1 \leq i \leq N$$

Desde una perspectiva de notación, un Modelos Ocultos de Markov (HMM) se escribe típicamente como: $\lambda = \{A, B, \pi\}$

3.3.2. Generación de Observaciones

- Para la detección de irrupciones se genera las observaciones
Eligiendo un estado inicial $q_1 = s_i$, basado en la distribución del estado inicial π .
- Para $t = 1, 2, \dots, T$:
 - Elegir $a_t = v_k$ en función de la distribución de probabilidad del símbolo en el estado s_i , $b_i(k)$.
 - Transición a un nuevo estado $q_{t+1} = s_j$ según la distribución de probabilidad de la transición de estado para el estado s_i , a_{ij} .
- Incrementar en 1, volver al paso 2 si $t \leq T$, de lo contrario, terminar

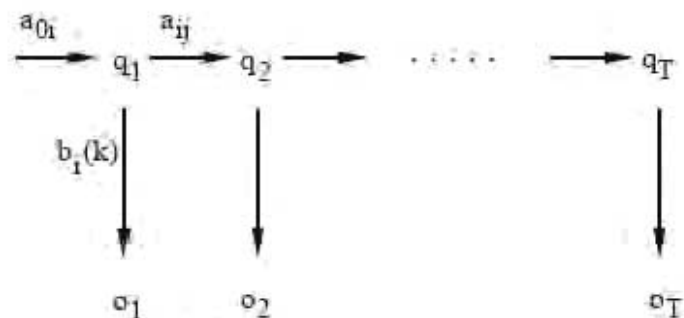


Figura N°15: Transiciones
Fuente: [Elaboración Propia]

3.3.3. Representación del diagrama de Estado



Figura N°16: Transiciones
Fuente: [Elaboración Propia]

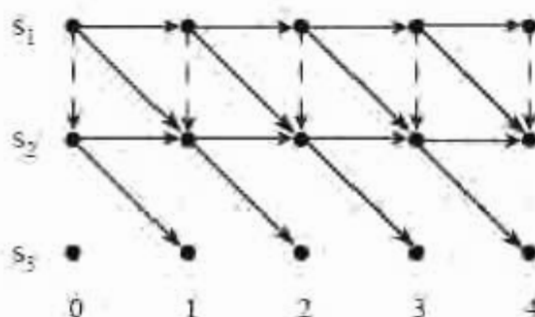


Figura N°17: Transiciones
Fuente: [Elaboración Propia]

La línea con guiones representa una transición nula, en la que no se genera ningún símbolo de observación de puertos.

3.3.4. Proceso de los Modelos Ocultos de Markov

A continuación se describe los problemas que detectan los Modelos Ocultos de Markov (HMM) y la solución que se plantea:

- Puntuación: Dada una secuencia de observación de puertos $O = \{o_1, o_2, \dots, o_T\}$ y un modelo $\lambda = \{A, B, \pi\}$, ¿cómo se calcula $P(O | \lambda)$, la probabilidad de la secuencia de observación?

Algoritmo de avance-retroceso

- Ajuste: Dada una secuencia de observación de puertos $O = \{o_1, o_2, \dots, o_T\}$, ¿cómo se elige una secuencia de estado $Q = \{q_1, q_2, \dots, q_T\}$ que de algún modo sea óptima?

Algoritmo de *Viterbi*

- Entrenamiento: ¿Cómo ajustamos los parámetros del modelo

$$\lambda = \{A, B, \pi\} \text{ Para maximizar } P(O | \lambda)$$

Algoritmo de reestimación de *Baum-Welch*

3.3.5. Cómputo de $P(O|\lambda)$

$$P(O|\lambda) = \sum P(O, Q|\lambda)$$

$$P(O, Q|\lambda) = P(O|Q, \lambda)P(Q|\lambda)$$

Considere la secuencia del estado fijo:

$$Q = \{q_1, q_2, \dots, q_T\}$$

$$P(O|Q, \lambda) = b_{q_1}(o_1)b_{q_2}(o_2) \dots b_{q_T}(o_T)$$

$$P(Q|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T}$$

$$P(O|Q, \lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T}$$

Por tanto:

$$P(O|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T}$$

3.3.6. El Algoritmo de Avance

Se define la variable de avance $\alpha_t(i)$, como la probabilidad de la secuencia de observación parcial hasta el tiempo t y estado s_i en el tiempo t , dado el modelo:

$$\alpha_t(i) = P(O_1 O_2 \dots O_t, q_t = s_i | \lambda)$$

Se puede demostrar fácilmente que:

$$\alpha_1(i) = \pi_i b_i(O_1), 1 \leq i \leq N$$

$$P(O|\lambda) = \alpha_T(i)$$

3.3.7. Ilustración del Algoritmo de avance

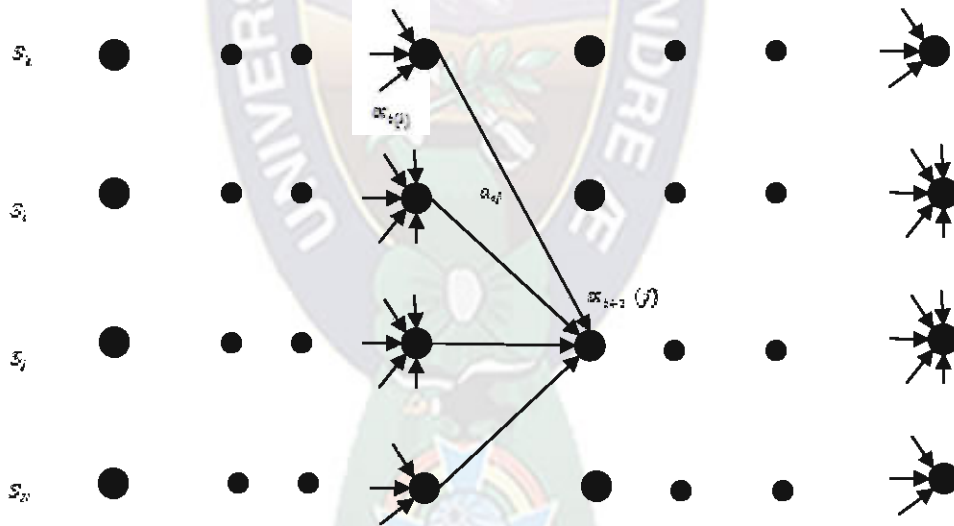


Figura N°18: Ilustración del algoritmo de Avance
Fuente: [Elaboración Propia]

3.3.8. Algoritmo de Retroceso

Del mismo modo se define las variables de retroceso $\beta_t(i)$, como la probabilidad de la secuencia de observaciones desde el tiempo $t + 1$, hasta el final dado el estado en el tiempo t y el modelo.

$$\beta_T(i) = P(O_{t+1} O \dots O_T | q_t = S_i, \lambda)$$

Puede demostrarse fácilmente que

$$\beta_t(i) = 1, \quad 1 \leq i \leq N$$

$$P(O/\lambda) = \sum_{i=1}^N \pi B(O_1) \beta_1(i)$$

Por inducción

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

$$t = T - 1, T - 2, \dots, 1$$

$$1 \leq i \leq N$$

3.3.9. Ilustración del procedimiento de Retroceso

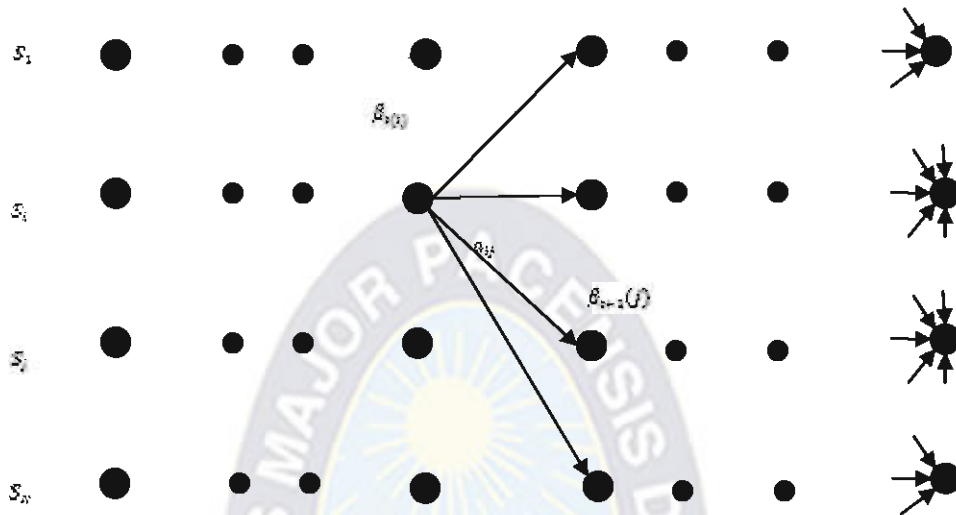


Figura N°19: Ilustración del procedimiento de Retroceso
Fuente: [Elaboración Propia]

3.3.10. Secuencias Óptimas de Estado

Un criterio selecciona estados que son individualmente los más probables

- Este maximiza el número esperado de estados correctos
- Se define $\gamma_t(i)$ como la probabilidad de estar en el estado, si en el tiempo t dada la secuencia de observaciones y el modelo.

$$\gamma_t(i) = P(q_t = s_i | O, \lambda)$$

$$\sum_{i=1}^N \gamma_t(i) \mathbf{1}(vt)$$

Luego el Estado individualmente probable q_t , en el tiempo t es:

$$q_t = \operatorname{argmax}_i \gamma_t(i) \quad 1 \leq t \leq T$$

$$1 \leq i \leq N$$

Observando, se puede demostrar que:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(Q|I, \lambda)}$$

- El criterio de optimalidad individual presenta el problema de que la secuencia del estado óptimo puede no obedecer a las restricciones de transición de estado.
- Otro criterio de optimalidad consiste en elegir la secuencia de estado que maximice $P(Q, O | \lambda)$, hallado mediante el algoritmo de Viterbi.
- $S_{t,i}$ es la probabilidad más alta a lo largo de una trayectoria simple en el tiempo t , que da cuenta de las primeras observaciones t

$$S_t(i) = \max P(q_1, q_2 \dots q_{t-1}, q_t = S_i, O_1, O_2, O_3, I | \lambda)$$

Por inducción

$$S_{t+1}(j) = [\max_i b_1(i), a_{ij}] b_j(O_{t+1})$$

Para recuperar la secuencia de estado, debemos seguir la pista de la secuencia de estado que proporciona la mejor trayectoria, en tiempo t , al estado de S_t .

3.3.11. Algoritmo de Viterbi

➤ Inicialización

$$S_t(i) = \pi_i b_i(O_1) \quad 1 \leq i \leq N$$

$$\psi_1(i) = 0$$

➤ Recursión

$$S_t(j) = \max [b_{t-1}(i)a_{ij}] \quad 2 \leq t \leq T$$

$$\psi_t(j) = \arg \max [b_{t-1}(i)a_{ij}] \quad 2 \leq t \leq T$$

➤ Terminación

$$p^* = \max [b_T(i)]$$

$$q^* = \arg \max [b_T(i)]$$

➤ Trayectoria inversa (Secuencia de estado)

$$q_t^* = \psi_{t+1}(q_{t+1}^*), t = T-1, T-2, \dots, 1$$

3.3.12. Algoritmo de Reestimación de *Baum-Welch*

- Definir $\varepsilon_t(i, j)$ como la probabilidad de estar en el estado S_i en el tiempo t y en el estado S_j en tiempo $t+1$ dado el modelo y la secuencia de observación

$$\varepsilon_t(i, j) = P(q_t = S_i, q_{t+1} = S_j \mid O, \lambda)$$

- Luego

$$\varepsilon_t(i, j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O \mid \lambda)}$$

$$\gamma_t(i) = \sum_{j=1}^N \varepsilon_t(i, j)$$

Sumando

$$\gamma_z(t) = \varepsilon_z(j)$$

Se obtiene:

$$\sum_{t=1}^{T-1} \gamma_z(t) = \text{Número esperado de transiciones desde } S_z$$

$$\sum_{t=1}^{T-1} \gamma_z(t, j) = \text{Número esperado de transiciones desde } S_z \text{ a } S_j$$

El procedimiento de reestimación se observa en la gráfica siguiente:

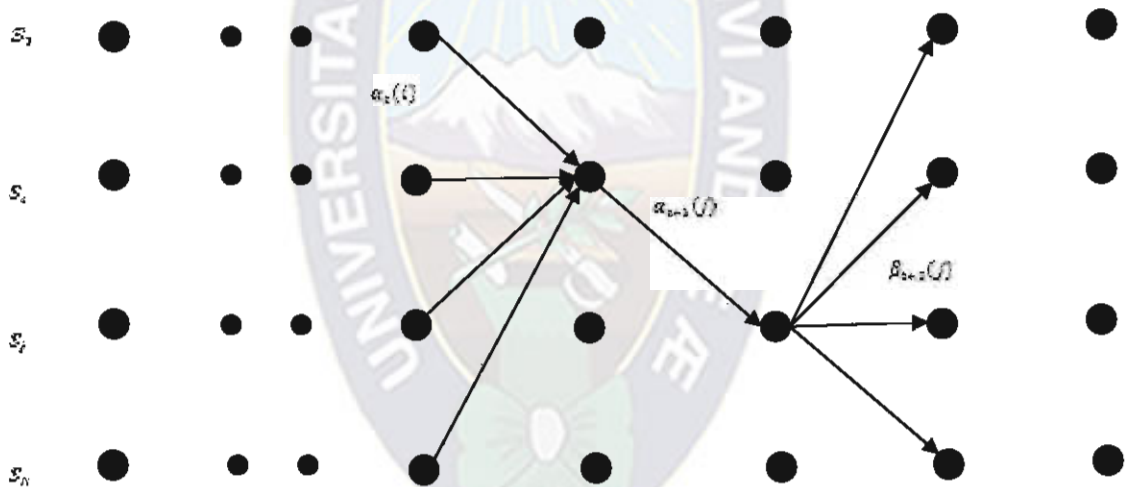


Figura N°20: Procedimiento de Reestimación
Fuente: [Elaboración Propia]

Y las fórmulas de restricción de *Baum – Welch* se interpretan:

$$\bar{\pi} = \text{Número esperado de tiempos en el estado } S_i \text{ el } t = 1 = \gamma_1(i)$$

$$\bar{\alpha}_{ij} = \frac{N^{\circ} \text{ de Transición desde el estado } S_i \text{ a } S_j}{N^{\circ} \text{ de Transición desde el estado } S_i}$$

$$\frac{\sum_{t=1}^{T-1} \varepsilon_t (j)}{\sum_{t=1}^{T-1} \alpha_t (i)}$$

$$b_j(k) = \frac{N^{\circ} \text{ Esp de tiempo en el Est } S_j \text{ a } S_i}{N^{\circ} \text{ de Esp de tiempo en el estado } S_j}$$

$$\frac{\sum_{t=1}^T \gamma_t (j)}{\sum_{t=1}^T \gamma_t (i)}$$

Si $\lambda = (A, B, \pi)$ Es el modelo Inicial

y $\bar{\lambda} = (\bar{B}, \bar{A}, \bar{\pi})$ El retroceso

Se puede demostrar que:

- *M inicial λ , define $\lambda = \lambda$*
- *$\bar{\lambda}$ es + probable que $\bar{\lambda}$*

3.4. PROTOTIPO

Mediante las siguientes imágenes se observa el comportamiento del prototipo de la detección de irrupciones en informática forense, donde también se utilizan los modelos ocultos de Markov. La siguiente figura muestra el interfaz de usuario inicial al ingresar al prototipo.



Figura N°21: Interfaz Inicial
Fuente: [Elaboración Propia]

Se observa los *links* de inicio, Forense, puertos,...Acerca de, Salir, que permiten ser mas entendible el prototipo.

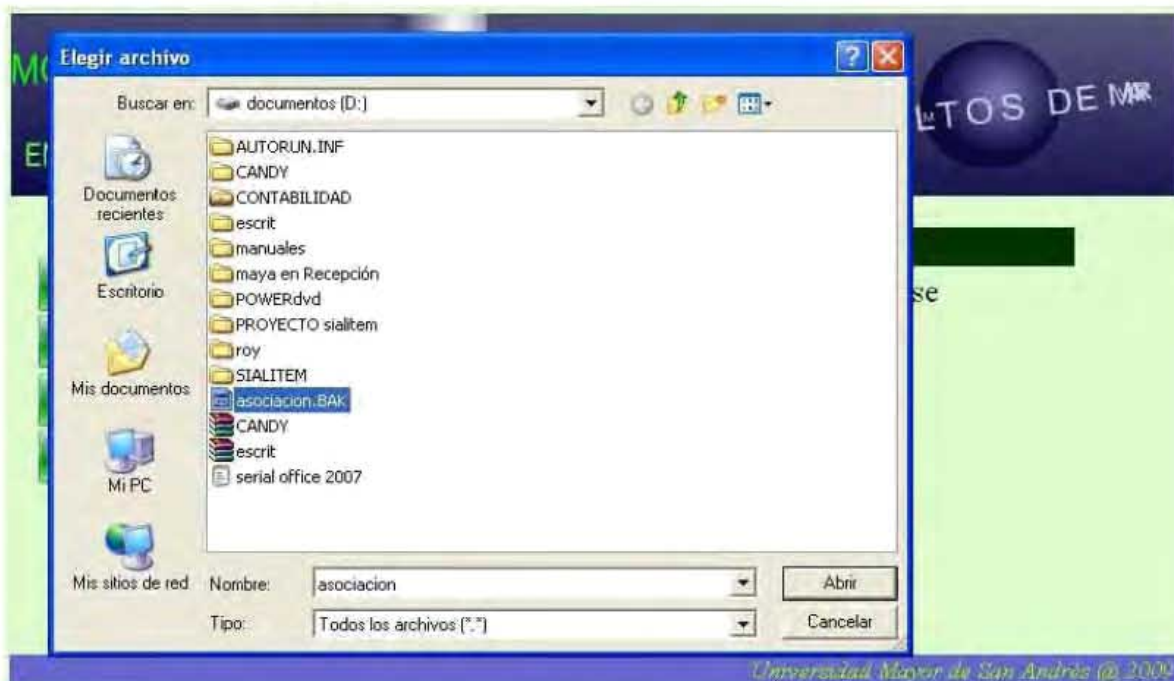


Figura N°22: Selección de documento para la detección de Irrupción
Fuente: [Elaboración Propia]

La opción "inicio", permite que vuelva a un estado inicial de ingreso al prototipo. La opción forense: permite verificar o analizar en el sistema que ha sido alterado o que se ha modificado y porque puerto existió la vulnerabilidad al equipo toma en cuenta todas los pasos que se han definido para la detección de irrupciones en informática forense.



Figura N°23: Archivo Seleccionado
Fuente: [Elaboración Propia]

La opción "puertos" muestra cómo los modelos ocultos de Markov, ayudan a detectar los puertos por donde ingresan las diferentes irrupciones hacia el equipo.

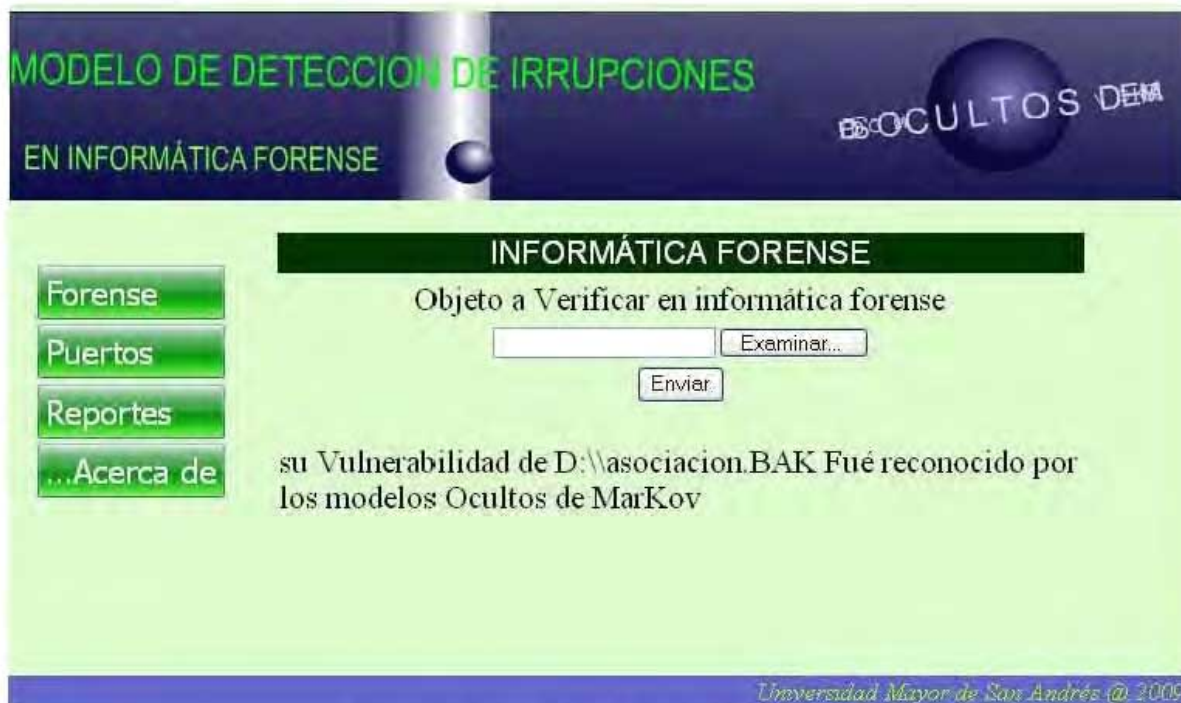


Figura N°24: Verificación de Vulnerabilidad
Fuente: [Elaboración Propia]

En la Figura N°25 se observa tres opciones: “Preservar Evidencia”, “Analizar evidencia” y “Generar evidencia” los cuales generan las figuras 26,27 y 28.

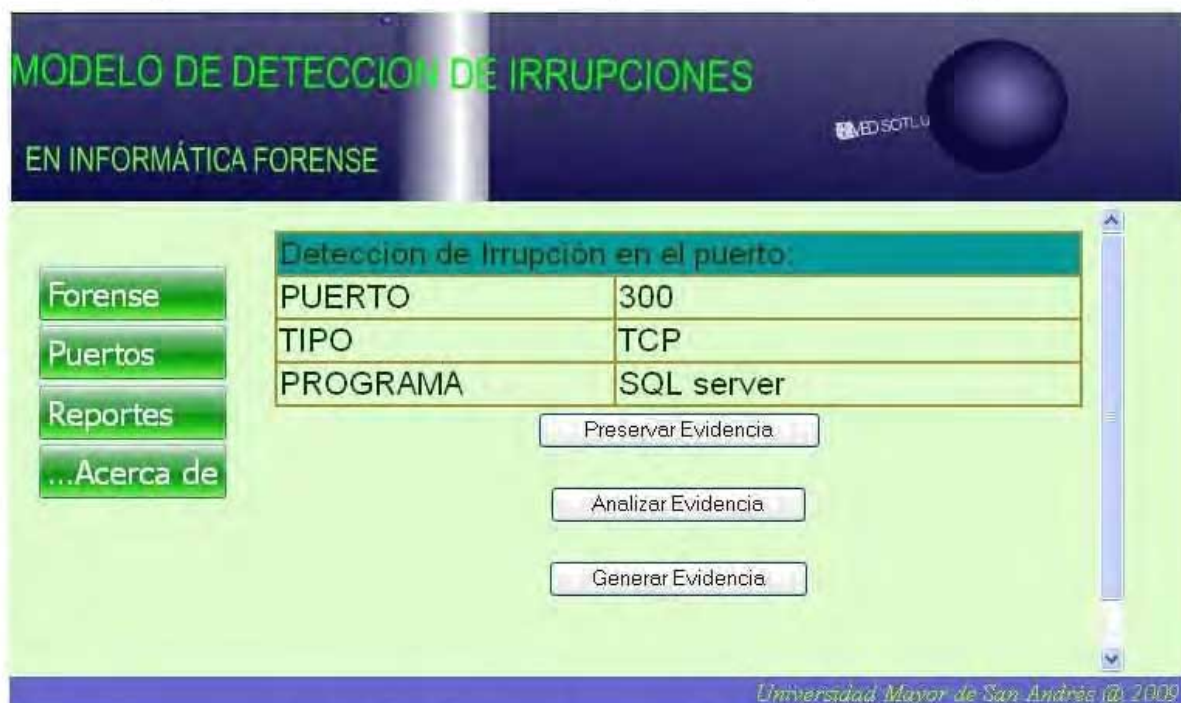


Figura N°25: Detección de Irrupción en el puerto
Fuente: [Elaboración Propia]

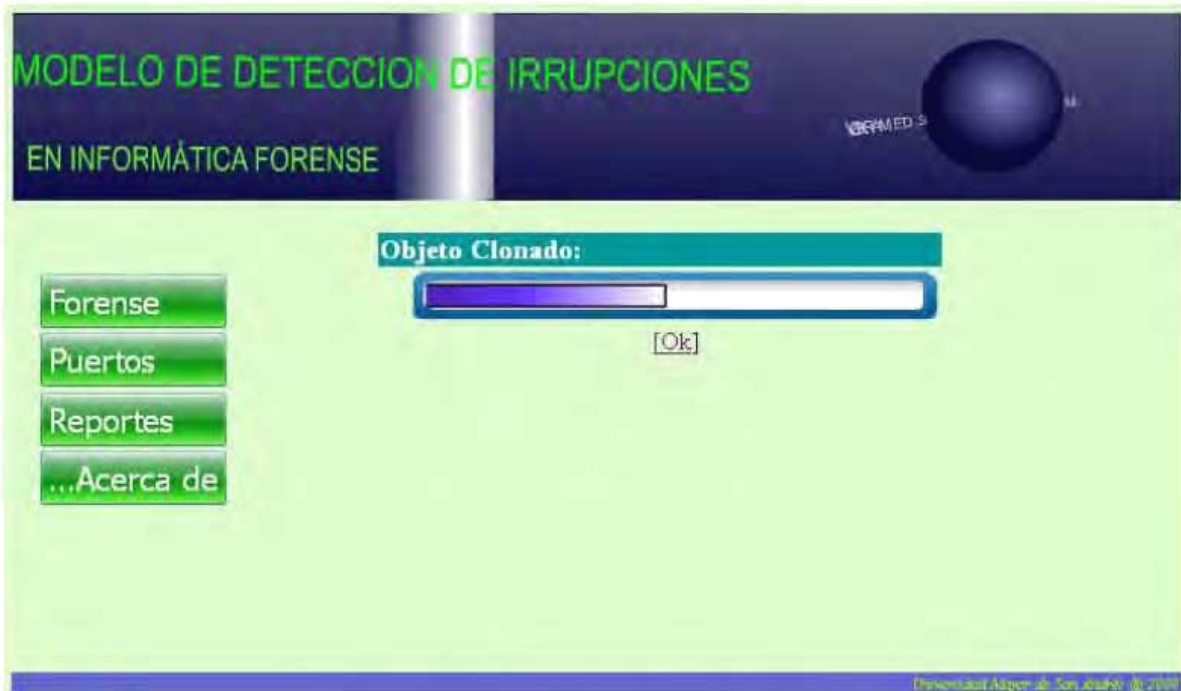


Figura N°26: Clonación de Archivo
Fuente: [Elaboración Propia]



Figura N°27: Analizando Evidencias
Fuente: [Elaboración Propia]



Figura N°28: Informe de Evidencia
Fuente: [Elaboración Propia]

En la Figura N° 29 muestra un reporte de las detecciones de irrupción apoyados con los modelos ocultos de Markov.



Figura N°29: Puertos detectados
Fuente: [Elaboración Propia]

3.5. PRUEBAS Y EXPERIMENTACIÓN

3.5.1. Definición De Variables De Prueba

La detección de irrupciones se hace conflictiva, pero las variables que permiten medir en el modelo planteado serán las que siguen a continuación:

Y: Detección de irrupciones en informática forense.

X_{i1} : Puertos detectados mediante los modelos Ocultos de Markov.

X_{i2} : Captura de la Evidencia sin modificación.

X_{i3} : Punto de Irrupción acertada.

La tabla siguiente define el comportamiento de las variables:

| | |
|----------|--|
| Y | La suma de Variables independientes acertadas en el reconocimiento Forense |
| X_{i1} | 0: No se detectó puerto |
| | 1: Si se detectó el puerto |
| X_{i2} | 0: No se ha capturado la evidencia sin modificación |
| | 1: Sí se ha capturado la evidencia sin modificación |
| X_{i3} | 0: se sabe cómo se irrumpió |
| | 1: No se Sabe cómo se irrumpió |

Tabla N° 2: Identificación de Variables.
Fuente: [Elaboración Propia]

3.5.2. Datos para la Experimentación

Se realizaron 50 pruebas de experimentación, de acuerdo a las mismas reportadas por el prototipo y el comportamiento del algoritmo de los modelos ocultos de Markov se observa lo siguiente:

| Nº | X ₁₁ | X ₁₂ | X ₁₃ | Y ₁ |
|----|-----------------|-----------------|-----------------|----------------|
| 1 | 1 | 1 | 0 | 2 |
| 2 | 1 | 1 | 1 | 3 |
| 3 | 1 | 1 | 1 | 3 |
| 4 | 0 | 1 | 1 | 2 |
| 5 | 1 | 0 | 1 | 2 |
| 6 | 1 | 0 | 1 | 2 |
| 7 | 1 | 1 | 1 | 3 |
| 8 | 0 | 1 | 1 | 2 |
| 9 | 1 | 1 | 1 | 3 |
| 10 | 1 | 1 | 0 | 2 |

Tabla N° 3: Pruebas para el experimento 1.
Fuente: [Elaboración Propia]

| Nº | X ₁₁ | X ₁₂ | X ₁₃ | Y ₁ |
|----|-----------------|-----------------|-----------------|----------------|
| 11 | 1 | 1 | 1 | 3 |
| 12 | 1 | 1 | 0 | 2 |
| 13 | 1 | 1 | 0 | 2 |
| 14 | 1 | 1 | 0 | 2 |
| 15 | 1 | 1 | 0 | 2 |
| 16 | 1 | 1 | 1 | 3 |
| 17 | 1 | 1 | 1 | 3 |
| 18 | 1 | 0 | 1 | 2 |
| 19 | 1 | 1 | 1 | 3 |
| 20 | 1 | 1 | 1 | 3 |

Tabla N° 4: Pruebas para el experimento 2.
Fuente: [Elaboración Propia]

| Nº | X ₁₁ | X ₁₂ | X ₁₃ | Y ₁ |
|----|-----------------|-----------------|-----------------|----------------|
| 21 | 0 | 1 | 1 | 3 |
| 22 | 1 | 1 | 1 | 3 |
| 23 | 1 | 1 | 1 | 3 |
| 24 | 1 | 1 | 1 | 3 |
| 25 | 1 | 1 | 0 | 2 |
| 26 | 1 | 1 | 1 | 3 |
| 27 | 1 | 1 | 1 | 3 |
| 28 | 1 | 1 | 1 | 3 |
| 29 | 1 | 1 | 1 | 3 |
| 30 | 0 | 1 | 1 | 2 |

Tabla N° 5: Pruebas para el experimento 3.
Fuente: [Elaboración Propia]

| Nº | X_{i1} | X_{i2} | X_{i3} | Y_i |
|----|----------|----------|----------|-------|
| 31 | 0 | 1 | 1 | 1 |
| 32 | 0 | 0 | 0 | 0 |
| 33 | 0 | 1 | 0 | 1 |
| 34 | 1 | 1 | 1 | 3 |
| 35 | 1 | 1 | 1 | 3 |
| 36 | 0 | 1 | 1 | 2 |
| 37 | 1 | 1 | 1 | 2 |
| 38 | 1 | 1 | 0 | 2 |
| 39 | 1 | 0 | 0 | 1 |
| 40 | 1 | 1 | 0 | 2 |

Tabla Nº 6: Pruebas para el experimento 4.
Fuente: [Elaboración Propia]

| Nº | X_{i1} | X_{i2} | X_{i3} | Y_i |
|----|----------|----------|----------|-------|
| 41 | 0 | 0 | 0 | 0 |
| 42 | 0 | 1 | 0 | 1 |
| 43 | 1 | 1 | 0 | 2 |
| 44 | 0 | 1 | 0 | 1 |
| 45 | 0 | 1 | 0 | 1 |
| 46 | 0 | 0 | 0 | 0 |
| 47 | 0 | 0 | 0 | 0 |
| 48 | 1 | 1 | 1 | 3 |
| 49 | 1 | 1 | 1 | 3 |
| 50 | 0 | 1 | 1 | 2 |

Tabla Nº 7: Pruebas para el experimento 5.
Fuente: [Elaboración Propia]

Por lo tanto, de acuerdo a los resultados se obtiene lo siguiente:

| Variables | Aciertos | Desaciertos |
|-----------|----------|-------------|
| X_{i1} | 35 | 15 |
| X_{i2} | 42 | 8 |
| X_{i3} | 31 | 19 |

Tabla Nº 8: Resultados de Prueba.
Fuente: [Elaboración Propia]

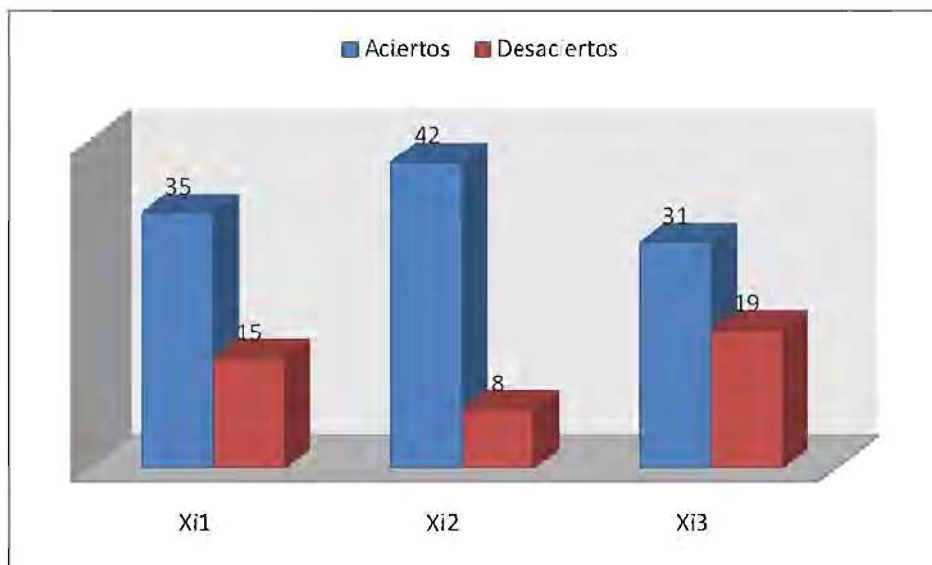


Figura N°30: Cuadro estadístico de Aciertos y desaciertos
Fuente: [Elaboración Propia]

| Yi | calificativo | % |
|-------|--------------|-----|
| 0 | 4 | 8 |
| 1 | 6 | 12 |
| 2 | 18 | 36 |
| 3 | 22 | 44 |
| Total | 50 | 100 |

Figura N°31: Porcentaje de Aciertos
Fuente: [Elaboración Propia]

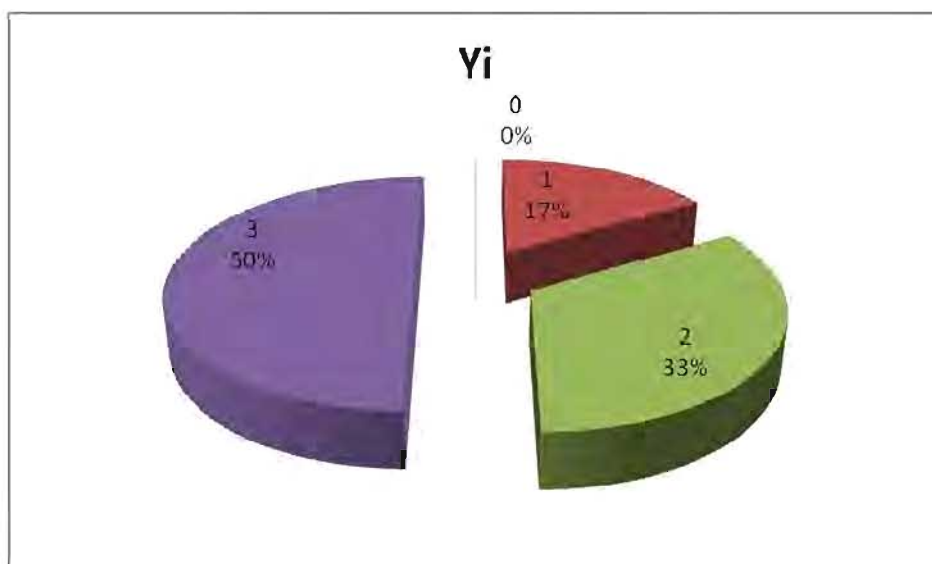


Figura N°32: Resultados en Porcentaje
Fuente: [Elaboración Propia]

3.5.3. Verificación de la Hipótesis

3.5.4. Resultados

$$Y_2+Y_3=50\%+33\%=83\%$$

De acuerdo al estudio realizado el porcentaje de acierto se mide en base a los porcentajes obtenidos, tomando la suma de ambas muestras a un porcentaje de acierto mayor al 80% en la detección de Irrupciones.



CAPÍTULO 4 CONCLUSIONES Y RECOMENDACIONES

4.1. CONCLUSIONES

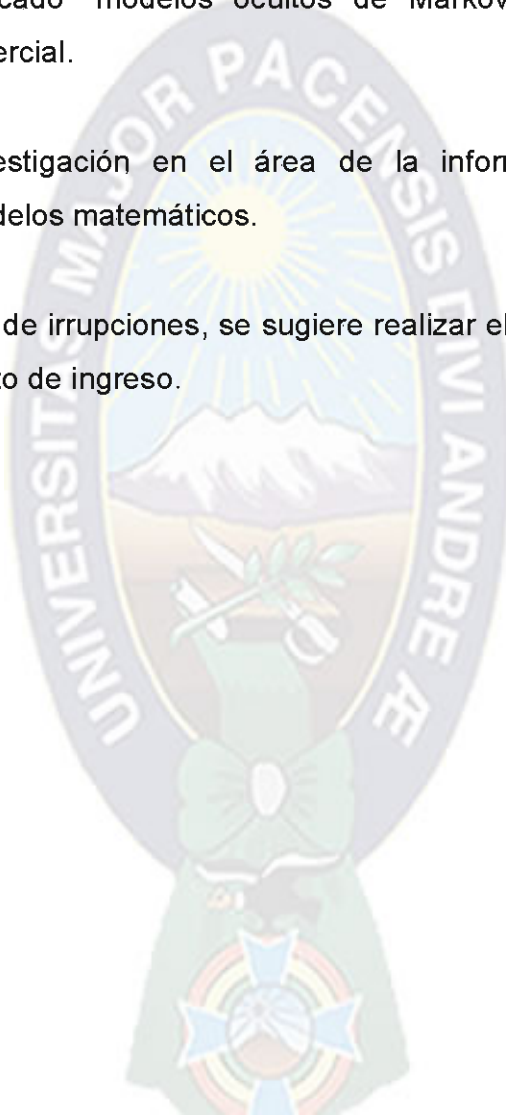
De acuerdo a la investigación realizada se concluye que:

- Se ha logrado obtener la información de los puertos mediante los modelos ocultos de Markov y se constituye la evidencia con base al mismo para llegar a la verdad sobre los hechos fraudulentos.
- Después de verificar la teoría acerca de la informática forense se ha llegado a desarrollar e investigar sobre los procedimientos para obtener las evidencias necesarias.
- Se ha logrado recoger información sobre manipulación informática y accesos no autorizados que ayude a resolver casos con gran dificultad.
- También se verifica que en Bolivia existe escasos estudios sobre informática forense y en la detección de irrupciones
- Finalmente, se puede conseguir fácilmente el *Software* y *Hardware* sin embargo la información relevante no.

4.2.RECOMENDACIONES

De acuerdo a las conclusiones realizadas y al proceso de investigación se recomienda que:

- El modelo aplicado “modelos ocultos de Markov”, sea aplicado en un *Software* comercial.
- Ampliar la investigación en el área de la informática forense por los diferentes modelos matemáticos.
- En la detección de irrupciones, se sugiere realizar el estudio separadamente por cada puerto de ingreso.



REFERENCIAS BIBLIOGRÁFICAS

BIBLIOGRAFÍA

- [ACT-2009] Actualidad TIC
Sistemas de Detección de Intrusos
Revista del instituto tecnológico de informática, 2009
- [ALL-1999] L. Allison,
Hidden Markov Models” School of Computer Science and
Software
Engineering, Monash University, Australia. 1999
- [CUR-2005] Curso Virtual
Procesos de Markov”
Univirtual. Universidad, Nacional de Colombia. 2005
- [HUA-2001] Huang, A. Acero y H. Hon,
Spoken Language Processing,
Prentice-Hall, 2001.
- [JEL-1997] Jelinek, T.
Statistical Methods for Speech Recognition.
MIT Press, 1997.
- [KLE-1975] Kleinrock, L.
Queuing Systems, Volume 1: Theory, Wiley Interscience
Publication, 1975.

- [MAT-2009] Matías Ison, Patricia Musulin, Laura Kamenetzky.
Modelos Ocultos de Markov”
biocomp, 2009
- [RAB-1993] L. Rabiner y B. Juang,
Fundamentals of Speech Recognition,
Prentice-Hall, 1993.
- [ROB-2007] Carrillo Aguilar, Roberto
design and manipulation of hidden markov models usingtk
tools. a tutorial
Coruña, 2007
- [ROB-2006] Robayo, L. Santana,
Modelo Matemático De Tráfico Para Detección De Intrusos En
Redes De Telecomunicaciones De Área Local Basado En
Modelos Ocultos De Markov –Estado Del Arte,
Universidad Nacional de Colombia, Seminario de Investigación
I, 2006.
- [SCA-2007] Scarfone , K. y Mell, P. “Guide to Intrusion Detection and
Prevention Systems (IDPS).
National Institute of Standards and technology, 2007.
- [TOS-2001] Tosum, U.
Hidden Markov models to Analyze user Behaviour in Network
Traffic”, Bilkent University, 2001

WEBGRAFÍA

- [ALA-2006] Alan Arehart, Ph. D.,
Hidden Markov Models (HMM)
<http://fulcrum.physbio.mssm.edu/~sdy/panningsearch/HiddenMarkovModels.htm>, 2006
- [AWE-2008] *AWeba*
Clasificación de Troyanos
<http://www.aWeba.com.ar/seguridad/clasificacion-de-troyanos.html>
2008
- [CHA-2003] Channelplanet
<http://www.channelplanet.com/?idcategoria=11833>
2003
- [MAS-2004] MasterMagazine
<http://www.mastermagazine.info/termino/5422.php>
2004
- [MVP-2009] Microsoft Most Valuable Professional
http://legalidadinformatica.blogspot.com/2009_07_01_archive.html, 2009.
- [WI1-2009] Wikipedia
http://es.wikipedia.org/wiki/Seguridad_inform%C3%A1tica
2009

- [WI2-2009] Wikipedia
http://es.wikipedia.org/wiki/Modelo_oculto_de_M%C3%A1rkov
2009
- [WI3-2009] Wikipedia
<http://es.wikipedia.org/wiki/Criptograf%C3%ADa>
2009
- [WI4-2009] Wikipedia
[http://es.wikipedia.org/wiki/Puerto_\(computaci%C3%B3n\)](http://es.wikipedia.org/wiki/Puerto_(computaci%C3%B3n))
2009
- [WI5-2009] Wikipedia
http://es.wikipedia.org/wiki/C%C3%B3mputo_forense
2009



ANEXO 1: NÚMEROS DE PUERTO

Números de puerto bien conocidos usados por TCP y UDP. También se añade algún otro puerto no asignado oficialmente por IANA, pero de interés general dado el uso extendido que le da alguna aplicación.

| Puerto/protocolo | Descripción |
|------------------|---|
| n/d / GRE | GRE (protocolo IP 47) Enrutamiento y acceso remoto |
| n/d / ESP | IPSec ESP (protocolo IP 50) Enrutamiento y acceso remoto |
| n/d / AH | IPSec AH (protocolo IP 51) Enrutamiento y acceso remoto |
| 1/tcp | Multiplexor TCP |
| 7/tcp | Protocolo <u>Echo</u> (Eco) Responde con eco a llamadas remotas |
| 7/udp | Protocolo <u>Echo</u> (Eco) Responde con eco a llamadas remotas |
| 9/tcp | Protocolo <u>Discard</u> Elimina cualquier dato que recibe |
| 9/udp | Protocolo <u>Discard</u> Elimina cualquier dato que recibe |
| 13/tcp | Protocolo <u>Daytime</u> Fecha y hora actuales |
| 17/tcp | <u>Quote of the Day</u> (Cita del Día) |
| 19/tcp | Protocolo <u>Chargen</u> Generador de caracteres |
| 19/udp | Protocolo <u>Chargen</u> Generador de caracteres |
| 20/tcp | <u>FTP</u> File Transfer Protocol (Protocolo de Transferencia de Ficheros) - datos |
| 21/tcp | <u>FTP</u> File Transfer Protocol (Protocolo de Transferencia de Ficheros) - control |
| 22/tcp | <u>SSH</u> , <u>scp</u> , <u>SFTP</u> |
| 23/tcp | <u>Telnet</u> comunicaciones de texto inseguras |
| 25/tcp | <u>SMTP</u> Simple Mail Transfer Protocol (Protocolo Simple de Transferencia de Correo) |
| 37/tcp | <u>time</u> |
| 43/tcp | <u>nickname</u> |
| 53/tcp | <u>DNS</u> Domain Name System (Sistema de Nombres de Dominio) |
| 53/udp | <u>DNS</u> Domain Name System (Sistema de Nombres de Dominio) |
| 67/udp | <u>BOOTP</u> BootStrap Protocol (Server), también usado por <u>DHCP</u> |

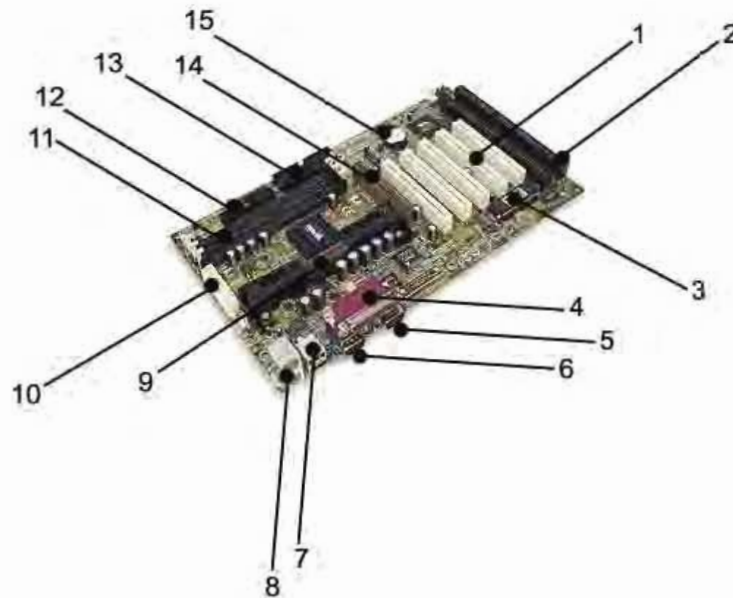
| | |
|---------|---|
| 68/udp | <u>BOOTP</u> BootStrap Protocol (Client), también usado por <u>DHCP</u> |
| 69/udp | <u>TFTP</u> Trivial File Transfer Protocol (Protocolo Trivial de Transferencia de Ficheros) |
| 70/tcp | <u>Gopher</u> |
| 79/tcp | <u>Finger</u> |
| 80/tcp | <u>HTTP</u> HyperText Transfer Protocol (Protocolo de Transferencia de HiperTexto) (<u>WWW</u>) |
| 88/tcp | <u>Kerberos</u> Agente de autenticación |
| 110/tcp | <u>POP3</u> Post Office Protocol (<u>E-mail</u>) |
| 111/tcp | <u>sunrpc</u> |
| 113/tcp | <u>ident</u> (auth) antiguo sistema de identificación |
| 119/tcp | <u>NNTP</u> usado en los grupos de noticias de <u>usenet</u> |
| 123/udp | <u>NTP</u> Protocolo de sincronización de tiempo |
| 123/tcp | <u>NTP</u> Protocolo de sincronización de tiempo |
| 135/tcp | <u>epmap</u> |
| 137/tcp | <u>NetBIOS</u> Servicio de nombres |
| 137/udp | <u>NetBIOS</u> Servicio de nombres |
| 138/tcp | <u>NetBIOS</u> Servicio de envío de datagramas |
| 138/udp | <u>NetBIOS</u> Servicio de envío de datagramas |
| 139/tcp | <u>NetBIOS</u> Servicio de sesiones |
| 139/udp | <u>NetBIOS</u> Servicio de sesiones |
| 143/tcp | <u>IMAP4</u> Internet Message Access Protocol (<u>E-mail</u>) |
| 161/tcp | <u>SNMP</u> Simple Network Management Protocol |
| 161/udp | <u>SNMP</u> Simple Network Management Protocol |
| 162/tcp | <u>SNMP-trap</u> |
| 162/udp | <u>SNMP-trap</u> |
| 177/tcp | <u>XDMCP</u> Protocolo de gestión de displays en <u>X11</u> |
| 177/udp | <u>XDMCP</u> Protocolo de gestión de displays en <u>X11</u> |
| 389/tcp | <u>LDAP</u> Protocolo de acceso ligero a Bases de Datos |
| 389/udp | <u>LDAP</u> Protocolo de acceso ligero a Bases de Datos |

| | |
|----------|--|
| 443/tcp | <u>HTTPS/SSL</u> usado para la transferencia segura de páginas <i>Web</i> |
| 445/tcp | Microsoft-DS (<u>Active Directory</u> , compartición en <u>Windows</u> , gusano <u>Sasser</u> , Agobot) |
| 445/udp | Microsoft-DS compartición de ficheros |
| 500/udp | <u>IPSec</u> ISAKMP, Autoridad de Seguridad Local |
| 512/tcp | <u>exec</u> |
| 513/tcp | <u>login</u> |
| 514/udp | <u>syslog</u> usado para logs del sistema |
| 520/udp | <u>RIP</u> |
| 591/tcp | <u>FileMaker</u> 6.0 (<i>alternativa para HTTP, ver puerto 80</i>) |
| 631/tcp | <u>CUPS</u> sistema de impresión de Unix |
| 666/tcp | identificación de <u>Doom</u> para jugar sobre TCP |
| 993/tcp | <u>IMAP4</u> sobre <u>SSL</u> (E-mail) |
| 995/tcp | <u>POP3</u> sobre <u>SSL</u> (E-mail) |
| 1080/tcp | <u>SOCKS</u> Proxy |
| 1337/tcp | suele usarse en máquinas comprometidas o infectadas |
| 1352/tcp | IBM Lotus Notes/Domino RCP |
| 1433/tcp | Microsoft-SQL-Server |
| 1434/tcp | Microsoft-SQL-Monitor |
| 1434/udp | Microsoft-SQL-Monitor |
| 1494/tcp | <u>Citrix MetaFrame</u> Cliente ICA |
| 1512/tcp | <u>WINS</u> |
| 1521/tcp | <u>Oracle</u> listener por defecto |
| 1701/udp | Enrutamiento y Acceso Remoto para VPN con L2TP. |
| 1723/tcp | Enrutamiento y Acceso Remoto para VPN con PPTP. |
| 1761/tcp | <u>Novell</u> Zenworks Remote Control utility |
| 1863/tcp | <u>MSN Messenger</u> |
| 1935/ | <u>FMS</u> Flash Media Server |
| 2049/tcp | <u>NFS</u> Archivos del sistema de red |
| 2082/tcp | <u>CPanel</u> puerto por defecto |

| | |
|----------|---|
| 2086/tcp | <u>Web Host Manager</u> puerto por defecto |
| 2427/udp | Cisco <u>MGCP</u> |
| 3030/tcp | <u>NetPanzer</u> |
| 3030/udp | <u>NetPanzer</u> |
| 3128/tcp | <u>HTTP</u> usado por <u>Web caches</u> y por defecto en <u>Squid cache</u> |
| 3128/tcp | <u>NDL-AAS</u> |
| 3306/tcp | <u>MySQL</u> sistema de gestión de bases de datos |
| 3389/tcp | RDP (<u>Remote Desktop Protocol</u>) |
| 3396/tcp | <u>Novell</u> agente de impresión NDPS |
| 3690/tcp | <u>Subversion</u> (sistema de control de versiones) |
| 4662/tcp | <u>eMule</u> (aplicación de compartición de ficheros) |
| 4672/udp | <u>eMule</u> (aplicación de compartición de ficheros) |
| 4899/tcp | RAdmin (<u>Remote Administrator</u>), herramienta de administración remota (normalmente <u>troyanos</u>) |
| 5000/tcp | <u>Universal plug-and-play</u> |
| 5060/udp | <u>Session Initiation Protocol</u> (SIP) |
| 5190/tcp | <u>AOL</u> y <u>AOL Instant Messenger</u> |
| 5222/tcp | <u>XMPP/Jabber</u> conexión de cliente |
| 5223/tcp | <u>XMPP/Jabber</u> puerto por defecto para conexiones de cliente SSL |
| 5269/tcp | <u>XMPP/Jabber</u> conexión de servidor |
| 5432/tcp | <u>PostgreSQL</u> sistema de gestión de bases de datos |
| 5517/tcp | <u>Setiqueue</u> proyecto SETI@Home |
| 5631/tcp | <u>PC-Anywhere</u> protocolo de escritorio remoto |
| 5632/udp | <u>PC-Anywhere</u> protocolo de escritorio remoto |
| 5400/tcp | <u>VNC</u> protocolo de escritorio remoto (usado sobre <u>HTTP</u>) |
| 5500/tcp | <u>VNC</u> protocolo de escritorio remoto (usado sobre <u>HTTP</u>) |
| 5600/tcp | <u>VNC</u> protocolo de escritorio remoto (usado sobre <u>HTTP</u>) |
| 5700/tcp | <u>VNC</u> protocolo de escritorio remoto (usado sobre <u>HTTP</u>) |
| 5800/tcp | <u>VNC</u> protocolo de escritorio remoto (usado sobre <u>HTTP</u>) |
| 5900/tcp | <u>VNC</u> protocolo de escritorio remoto (conexión normal) |

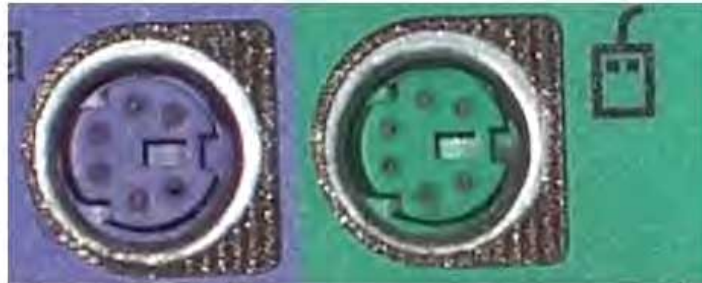
| | |
|-----------|--|
| 6000/tcp | <u>X11</u> usado para X-windows |
| 6112/udp | <u>Blizzard</u> |
| 6129/tcp | <u>Dameware Software</u> conexión remota |
| 6346/tcp | <u>Gnutella</u> compartición de ficheros (Limewire, etc.) |
| 6347/udp | <u>Gnutella</u> |
| 6348/udp | <u>Gnutella</u> |
| 6349/udp | <u>Gnutella</u> |
| 6350/udp | <u>Gnutella</u> |
| 6355/udp | <u>Gnutella</u> |
| 6667/tcp | <u>IRC IRCU</u> <i>Internet</i> Relay Chat |
| 6881/tcp | <u>BitTorrent</u> puerto por defecto |
| 6969/tcp | <u>BitTorrent</u> puerto de tracker |
| 7100/tcp | Servidor de Fuentes <u>X11</u> |
| 7100/udp | Servidor de Fuentes <u>X11</u> |
| 8000/tcp | <u>iRDMI</u> por lo general, usado erróneamente en sustitución de 8080. También utilizado en el servidor de streaming ShoutCast. |
| 8080/tcp | <u>HTTP HTTP-ALT</u> ver puerto 80. <u>Tomcat</u> lo usa como puerto por defecto. |
| 8118/tcp | <u>privoxy</u> |
| 9009/tcp | <u>Pichat</u> peer-to-peer chat server |
| 9898/tcp | Gusano Dabber (troyano/virus) |
| 10000/tcp | <u>Webmin</u> (Administración remota <i>Web</i>) |
| 19226/tcp | <u>Panda Security</u> Puerto de comunicaciones de Panda Agent. |
| 12345/tcp | <u>NetBus en:NetBus</u> (troyano/virus) |
| 31337/tcp | <u>Back Orifice</u> herramienta de administración remota (por lo general troyanos) |



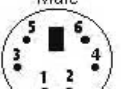
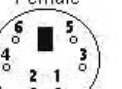
ANEXOS 2: TARJETA MADRE Y SUS PUERTOS



- Conectores PCI
- Conectores ISA
- Conectores de Sonido
- Puerto Paralelo
- Conectores de Mouse Serial
- Conectores de Joytick
- Conectores USB
- Conectores PS2 de teclado y mouse
- Microprocesador
- Conectores de Floppy Disk
- Conectores para Memorias RAM
- Conectores de disco Master
- Conectores de disco Slave
- Conectores AGP
- Bateria de la tarjeta madre

ANEXO 3: PUERTOS PS/2



| | | |
|---|---|---|
| <p>Male</p>  <p>(Plug)</p> | <p>Female</p>  <p>(Socket)</p> | <p>5-pin DIN (AT/XT):</p> <ul style="list-style-type: none"> 1 - Clock 2 - Data 3 - Not Implemented 4 - Ground 5 - Vcc (+5V) |
| <p>Male</p>  <p>(Plug)</p> | <p>Female</p>  <p>(Socket)</p> | <p>6-pin Mini-DIN (PS/2):</p> <ul style="list-style-type: none"> 1 - Data 2 - Not Implemented 3 - Ground 4 - Vcc (+5V) 5 - Clock 6 - Not Implemented |

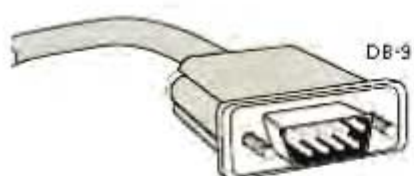
ANEXO 4: PUERTOS USB



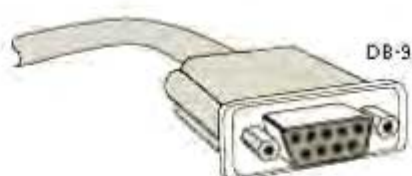
ANEXO 5: PUERTOS SERIALES Y PARALELOS

PUERTOS SERIALES

Macho



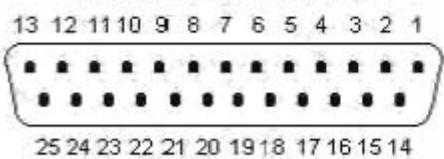
Hembra



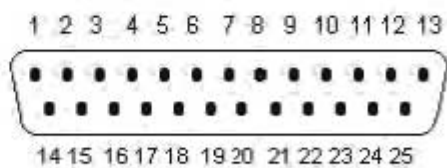
PUERTOS PARALELOS



Conector DB25 hembra del PC



Conector macho del Centronic al PC

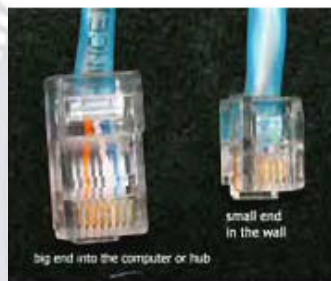


ANEXO 6: PUERTOS RJ Y VGA

PUERTOS RJ-11



PUERTOS RJ-45



PUERTOS VGA



CONECTOR VGA



ANEXO 7: RCA

PUERTOS RCA



ANEXO 8: GPS

GPS, Sistema de Posicionamiento Global



El GPS (Global Position System / Sistema de Posicionamiento Global), es un GNSS (Sistema Global de Navegación por Satélite), que permite determinar absolutamente en todo el mundo la posición de un objeto determinado, un vehículo o incluso una persona, con una precisión hasta de pocos metros.

Fue desarrollado e instalado por el Departamento de Defensa de los Estados Unidos, aunque originalmente se ha pretendido identificar a su creación a los gobiernos belga y franceses.

Funciona mediante una red de 27 satélites (24 operativos y 3 de respaldo), en órbita sobre el globo a 20.200 km con trayectorias sincronizadas para cubrir toda la superficie de la Tierra.

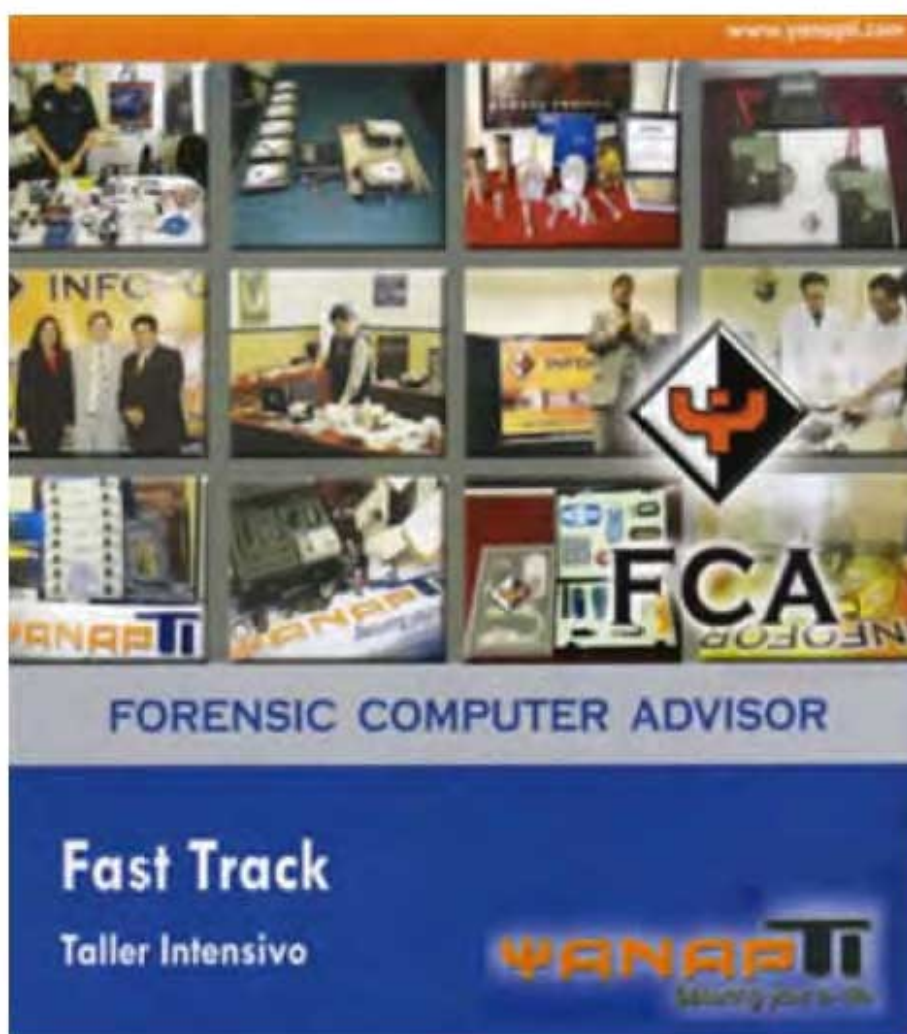
Para determinar la posición, el aparato que se utiliza para ello localiza automáticamente como mínimo cuatro satélites de la red, de los que recibe unas señales indicando la posición y el reloj de cada uno de ellos.

Por ello, se sincroniza el reloj del GPS calculando luego el retraso de las señales. Finalmente, por triangulación calcula la posición en que éste se encuentra.

ANEXO 9: YANAPTI

Precursores de la informática forense en Bolivia

ΨANAPTI
Securing your e-life



www.yanapti.com

INFOFOR

Ψ

FCA

INFOFOR

FORENSIC COMPUTER ADVISOR

Fast Track
Taller Intensivo

ΨANAPTI
Securing your e-life