

Hacia una Taxonomía para Sistemas de Detección de Intrusos, Enfocada desde Minería de Datos

Pedro P. Pinacho D.

Universidad de Santiago de Chile, Facultad de Ingeniería, DIINF,GITS** Av.
Ecuador 3659 Santiago, Chile.

Resumen En este artículo se realiza una sencilla revisión a un conjunto de Sistemas de Detección de Intrusos, tanto comerciales como en proyectos de investigación o software Libre. Estos son inicialmente clasificados de forma elemental y correlacionados con una taxonomía de Minería de datos para su análisis, tratando de establecer un marco de referencia común para futuros estudios.

Palabras Clave: *taxonomía, IDS, survey, Minería de Datos*

1. Introducción

Los enfoques tradicionales para enfrentar la amenaza de las intrusiones, es decir, el momento donde las políticas de seguridad y dispositivos de autenticación han sido traspasados. Son en general poco prácticos, y se basan en la mayoría de los casos en soluciones ad-hoc, lentas, costosas y erráticas. Por otro lado los actuales IDSs¹, tienen extensibilidad limitada, como pobre adaptabilidad a la incorporación de detección de nuevos tipos de ataque.

Entendiendo una intrusión como *Cualquier acción que atente o comprometa la integridad, confidencialidad, o disponibilidad de algún recurso*[9]. La detección y tratamiento de estas situaciones, puede ser vista como la última línea de defensa, la cual actúa en el momento de la transgresión o posterior a ella[8].

En la práctica las manifestaciones del comportamiento del usuario en un sistema son múltiples, y los datos desprendidos de tales transacciones es por volumen inteligible a escala de análisis humano, al momento de querer determinar posibles transgresiones. Es por esto que las técnicas de *minería de datos*, se ajustan a este tipo de aplicaciones.

Por otro lado la aparición de nuevos problemas de seguridad es continua, lo cual se denota en el gran volumen de tráfico que presentan grupos de discusión y listas de correo como *CERT*, *bugtraq*, *alt2600* y *Phrack*. Siendo uno de los enfoques más tradicionales es el de corregir cada posible error de diseño que lleve

** Grupo de Investigación de Tópicos de Seguridad (<http://gits.diinf.usach.cl>)

¹ Intrusion Detection Systems

al atacante a tener la posibilidad de generar algún *exploit*², lo cual es excesivamente caro y cercanamente imposible. Justificándose aún más la existencia de una última línea de defensa, cuando se sabe que:

- El 80 % de los ataques que sufre una entidad son internos.
- El hurto de información a aumentado un 250 %, en los últimos 5 años.
- El 99 % de las más grandes compañías reportan al menos un incidente mayor de seguridad informática.
- El fraude en telecomunicaciones alcanza ya los diez billones de dólares solo en EEUU.[12]

Entre los beneficios de los sistemas de detección de intrusos podemos distinguir:

- Protección en tiempo real, lo cual nos permite reportar los incidentes de seguridad y tomar las acciones pertinentes.
- Permitir la identificación de la falla, con lo cual los encargados de las seguridad puedan identificarla y corregirla.

Cualquiera sea el tipo de IDS, se asume que las actividades de los usuarios pueden ser monitoreadas, ya sean estas intrusivas o no. Lo que si diferentes sistemas de detección usan diferentes características.

2. Clasificación de IDS Según el Enfoque del Problema

Es de conceso en el área, que existen dos enfoques principales para la detección de intrusos.

Detección de mal uso (firmas): El cual se basa en la detección de ataques bien conocidos, los cuales han sido representados como patrones, que se correlacionan con las actividades de los usuarios del sistema o recursos.

Detección de Anomalía: En este caso, el comportamiento normal del usuario o recurso es caracterizado a través de perfiles de comportamiento de estos, los cuales son usados para detectar intrusiones a través de las desviaciones que se presenten.

2.1. Detección de Anomalía

La detección de anomalía consiste en establecer patrones normales de comportamiento para usuarios, programas u otro recurso de interés en el sistema, detectando a través de las desviaciones las posibles intrusiones en el sistema. Existen múltiples aproximaciones de carácter estadístico como SRI's IDDES, NIDES, en los cuales el comportamiento de los usuarios consiste en un conjunto de mediciones estadísticas, enfoque muchas veces llamado *vector de características*.

Por ejemplo en NIDES las mediciones son del siguiente tipo:

² Ataque basado en una vulnerabilidad del Sistema

Medición Ordinal: Este es un conteo de algunos cuantificables aspectos de comportamiento observables (Ej. tiempo de CPU usado).

Medición Categoría: Se clasifica el comportamiento sobre un conjunto finito de categorías, estas pueden ser:

Medición Binaria: en este caso la categoría esta presente o no, usada muchas veces para categorizar eventos infrecuentes.

Medición Lineal: En este caso indica las veces que ocurrió cada categoría en el comportamiento observado.

Estos factores se usan con marcas de peso para ponderarlos y de esta forma determinar el grado de anormalidad. Los perfiles de usuarios también son actualizados periódicamente, para de esta forma incorporar el comportamiento nuevo normal de los usuarios supervisados.

Una de las principales características de la detección de anomalía, es que el enfoque puede servir para detectar amenazas no conocidas, por el hecho de que no requiere información a priori para establecer una reacción en contra de una variación significativa, respecto a los patrones codificados en el perfil.

Por otro lado este enfoque presenta las siguientes complejidades:

- La selección de las características a medir del sistema pueden variar grandemente dependiendo del ambiente de computo.
- La parametrización correcta de los umbrales de desviación es ad-hoc.
- El comportamiento del usuario suele cambiar dinámicamente, lo cual puede volver inconsistentes los datos del perfil de comportamiento, lo cual podría aumentar al mediano plazo las tasa de falsos positivos.
- Considerando los sistemas basados en vectores de características, algunas intrusiones solo pueden ser detectadas mediante una investigación de las correlaciones entre una secuencia de eventos, figurando de forma normal en las estadísticas.
- Los sistemas basados en estadística pueden ser entrenados, por un intruso bien informado para que este acepte su comportamiento intrusivo como normal.

Entre los sistemas revisados y líneas de investigación, los siguientes proveen capacidades de detección de anomalía: MADAM ID, Línea de T. Y. Lin, NIDES, Línea de Stephanie Forrest, IDES, Haystack, MIDAS, W&S, ComputerWatch, NSM, Hyperview, DPEM, Janus, Esmerald, Bro.

2.2. Detección de Mal Uso

La detección de mal uso, pretende almacenar patrones específicos de actividades intrusivas en los sistemas computacionales, monitoreando en búsqueda de la aparición de estos en los recursos o usuarios monitoreados y reportando las correlaciones. Existen varias aproximaciones en la detección de mal uso, las cuales difieren en representación, como en los algoritmos de correlación utilizados.

Por ejemplo el Sistema NIDES, también posee características para la detección de mal uso a través de un componente de *sistema experto basado en*

reglas, un ejemplo de reglas para esto es: *más de tres intentos infructuosos de entrada al sistema, es un intento de penetración en el sistema*. Otro ejemplo es STAT, el cual usa análisis de estado de transición.

En la práctica la mayoría de los IDS se clasifican dentro de esta categoría, especialmente los de uso comercial actual, como lo son CISCO IDS 4235, Intrusion SecureNet 7145C, StealthWatch M100, NID 200, Nokia IP 530, OneSecure, Recourse ManHunt y el OSS Snort y otros como MADAM ID, T Y. Lin, NIDES, IDIOT, Haystack, MIDAS, NADIR, USTAT, JiNao, ESMEERALD, Bro, Rosset, DIDS, ASAX

La principal ventaja del enfoque es que puede eficientemente detectar e identificar ataques en proceso, pero no es bueno para detectar amenazas desconocidas, los inconvenientes apuntan a que los patrones de ataque requieren una actualización sistemática de muy alto costo.

Varios de los IDS actuales usan de forma conjunta los enfoques de detección de anomalía y el de detección de mal uso, radicando el problema muchas veces en la desconexión casi completa entre ambas partes.

Existen algunas conceptualizaciones que tratan de determinar el ataque a través de un sistema de detección de mal uso, mientras determinan las posibles variantes del ataque mediante la detección de anomalía basada en el comportamiento normal del ataque.

3. Clasificación de IDS, Según Origen de Datos

3.1. IDS basados en Host

Otra forma de clasificar los IDS, se según el tipo de datos que recopilan en su monitoreo, los IDS basados en host utilizan para la detección de intrusiones, información proveniente de un host particular, al cual se protege, esto es, información acerca de sesiones establecidas, procesos ejecutados, uso de recursos de la máquina monitoreada, y revisión del comportamiento del usuario sobre esta (GUI, consola..). Este es el caso de COAST y Pinacho[1][2][3][14][15], y cualquier IDS basado en general de eventos (*logs*), los cuales pueden ser incluso sobre las actividades de la red, pero según el host.

3.2. IDS basados en Red

Otra variante de IDS, tiene un planteamiento más especializado, esto es detección de intrusos en la red, para lo cual utilizan técnicas específicas para tal dominio, como detección de anomalías en protocolos conocidos, o búsqueda de firmas de ataque distintivos, además de establecimiento de formas normales de tráfico en un segmento de red, usualmente interactúan con la red de forma pasiva, escuchando en modo promiscuo los paquetes que circulan en un segmento. Estos usualmente presentan una portabilidad mayor, debido a que se basan en elementos más estándar (la red), sin recurrir a recursos más específicos de los sistemas operativos como lo son los logs.

Problemas de Sistemas de Detección de Intrusos Actuales Una forma de medir los IDS son a través de su *efectividad, adaptabilidad y extensibilidad*[9]. Un IDS es efectivo, si presenta un alto grado de clasificación correcta entre la conducta intrusiva y la normal, esto es, altos grados de verdaderos positivos y baja cantidad de falsos positivos. La adaptabilidad tiene que ver, con la detección de ataques con modificaciones o bien, poder aceptar las variaciones normales del comportamiento de los usuarios.[14]

La efectividad de los IDS usualmente se ve diezmada por el hecho de que las reglas son introducidas manualmente, supeditando las capacidades del sistema al conocimiento experto del administrador responsable.

Por otro lado la adaptabilidad ante nuevas vulnerabilidades es lenta, puesto que las curvas de aprendizaje existentes respecto a la razón de detección de nuevas falencias de seguridad.

La extensibilidad y configuración en IDS también es compleja, debido a que están conceptualizados para ambientes de computo determinados y además suelen poseer estructuras monolíticas de difícil mantención.

En la práctica en la valoración de los actuales IDS, incluso los comerciales, deben considerarse medidas adicionales para revisar su rendimiento, debido a que expuestos a ambientes los suficientemente hostiles fallan incluso colapsan[13], debido principalmente a sobrecargas de bases de datos sumado a un notable número de falsos positivos.

3.3. Clasificación de la Minería de Datos Aplicada en Seguridad

Algunos autores afirman que la minería de datos, es una área de las Ciencias de la Computación[26], y proveen distintas formas de clasificar las técnicas de minería de datos, no utilizando un criterio único de clasificación sino varios, como en el caso siguiente:

- Según la forma de obtener la información, a partir de datos históricos a utilizar.
- Tomando en cuenta la interacción entre el usuario y la máquina durante el proceso.
- Considerando el objetivo del proceso de minería de datos.

Respecto a la forma de obtener la información, las aplicaciones de detección de intrusos típicamente recurren a la agrupación, o particionamiento del comportamiento por sus características, por sobre la asociación de variables. Ahora si tomamos en cuenta la interacción entre la máquina y el usuarios, tenemos visiones parecidas para el caso de la detección de la anomalía y del mal uso, puesto que predomina el *aprendizaje supervisado*, debido a que las actividades de generación de perfil y correlación necesitan la intervención del usuario del sistema, para la determinación de conjuntos de entrenamientos, (caso anomalía). En el caso de la detección de mal uso, la acción del usuario es continua y necesaria para el ingreso de nuevos patrones de ataque al sistema. Finalmente respecto al objetivo, en el caso de la detección de intrusos, este es la clasificación, puesto

que todo el proceso se realiza para generar la distinción entre comportamiento intrusivo y comportamiento normal.

Las técnicas de minería de datos pueden ser clasificadas con independencia de los objetivos[26], debido que en muchos caso pueden ser utilizadas tanto como para predicción de comportamiento como para clasificación. En general las técnicas que son utilizadas en los sistemas para la detección de intrusos son las siguientes:

Aprendizaje basado en casos Corresponde a técnicas basadas en la búsqueda de patrones en los k vecinos más cercanos. Esto es, básicamente un método de búsqueda, pero útil para determinar grupos y distancias.

Estas técnicas son utilizadas para detección de anomalías a través de la transformación de secuencias temporales de datos en observaciones en un espacio métrico, que codifican dependencia entre los atributos[1]

Dentro de esta categoría quizás podemos clasificar algunas técnicas estadísticas basadas en vectores de características como lo es en la parte de detección de anomalía que encontramos en IDES y su sucesor NIDES. como también las utilizadas en detección de anomalía en Haystack[11] y MIDAS, MSN y JiNao, ESMERALD, HyperView y en general en prácticamente cualquier sistema de detección de anomalía.

Agrupaciones Difusas Permite determinar pertenencia en grados no estrictos. Usadas para la detección de anomalía, a través de la extracción de reglas de corrientes de datos continuas, siendo la ventaja de estas reglas sobre las exactas, el hecho que une la constante actualización necesaria, el costo de actualizar reglas exactas es demasiado alto, este tipo de técnicas suelen ser usadas para la detección de fraudes en telecomunicaciones.

Árboles de decisión Consiste en asumir bajo ciertos predicados, el comportamiento de los atributos que presentan los datos que se analizan. Permite construir modelos de clasificación.

Existen esquemas de reglas para alimentar base de datos que son parametrizados por el comportamiento de usuarios, a través de la generación de perfiles, o usándose estos para la detección de mal uso con reglas generadas por expertos, pero sensibilizadas por factores obtenidos del propio sistema, como es el caso de NADIR.[11], otro sistema basado en reglas es ASAX

Dentro de las técnicas de este tipo tenemos algoritmos de aprendizaje de regla como RIPPER[27], producidas por MADAM ID[9], utilizadas para clasificación por detección de anomalía.

Además algoritmos como C4.5 con algunas modificaciones, son utilizados para la generación automática de reglas que caractericen fraudes, es decir, mal uso en sistema de telecomunicaciones.[20]

Otros enfoques como W&S[11], utilizan e su esquema de detección de anomalías Sistemas Expertos basado en reglas generadas y actualizadas según los datos recolectados.

Descubrimiento de Reglas de Asociación Representan combinaciones de items o productos que contienen una condición y un resultado. Las reglas de asociación junto con los episodios frecuentes también son la base de algunos trabajos en el área[17]

Análisis de series de tiempo Esta técnica consiste en detectar cuando se producen secuencias de patrones entre transacciones para un determinado período de tiempo.

Este tipo de minería es de amplia utilización en las soluciones para detección de intrusos, es así que las encontramos en las comparaciones de similitud entre patrones de eventos o texto en el tiempo, como en el caso de trabajos basados en motivaciones biológicas [14] [21], basados en la búsqueda de divergencias ante patrones conocidos en perfiles de comportamiento normal, ya sea en el estudio de llamadas de sistemas de usuarios o bien en comandos de consola, como es el caso también caso de los trabajos de laboratorio COAST[1].

Existen también otras técnicas menos tradicionales como el uso de *Redes de Petri Coloreadas* para hacer pattern matching de patrones en secuencia, como en el caso de la técnica elegida por el laboratorio COAST con IDIOT. Otra forma de hacer descripciones de ataque o *firmas*, los cuales son un conjunto de operaciones ordenadas en el tiempo es a través de diagramas de estado de transición, como en el caso de los utilizados por USTAT[11], el problema radica en su carácter estático, por lo cual no clasifican dentro de la caracterización de minería de datos, caso parecido a DPEM.

Además existen otros tipos de técnicas ajenas a minería de datos, utilizadas en los IDSs como lo son los sistemas expertos tradicionales, los cuales por tener una estructura estática programada a priori y no establecida por la naturaleza de los datos auditados, caen fuera de la categoría de la MD, este es el caso del componente de detección de firmas de ataque de NIDES. Otro ejemplo es la conceptualización que hace MIDAS en la detección de ataques por firma, mediante un Sistema Experto, cuya única forma de actualización es manual. Otro ejemplo de esto es DIDS[11]

Otros sistemas no clasificables como Minería de datos, solo generan actividades de monitoreo y reporte, donde se pueden generar algunas alarmas, solo determinadas por la parametrización del SSO³, como es el caso de ComputerWatch[11]

También quedan fuera de esta clasificación sistemas basados en reglas, pero de tipo estático, como es el caso del componente de detección de firmas de ES-MERALD, o el sistema basado en políticas comercial Bro, como también los recientemente evaluados con pésimo resultado, como lo son Cisco IDS 4235[13][28], Intrusion SecureNet[30] 4145C, StealtWatch M100, NID200, Nokia IP 530[31], OneSecure[32], ManHunt[33] en su funciones de detección por firma y Snort[29].

4. Clasificación de los IDS Revisados

Basado en los clases entregadas, es que a continuación se entrega una clasificación sencilla de un número importante de proyectos de IDS, tanto libres como comerciales, de forma de establecer un marco de comparación común para tales herramientas.

³ Jefe de Seguridad

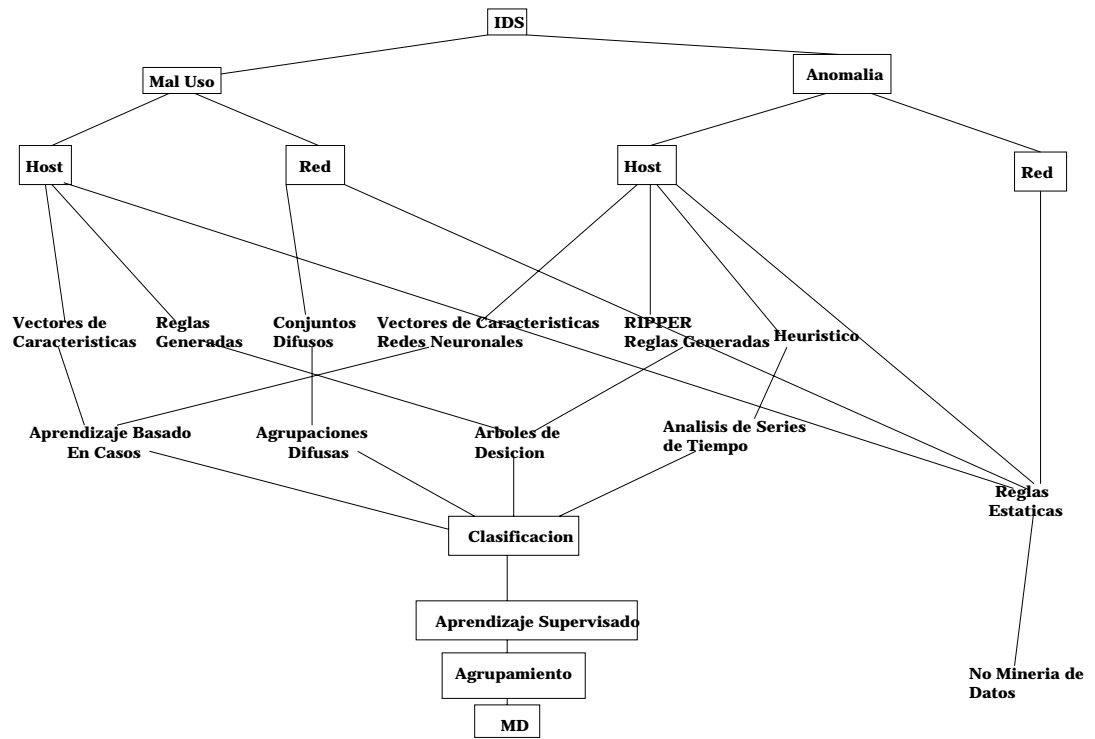


Figura 1. Clasificación de IDS desde MD

La clasificación generada y mostrada en Figura 1, entrega un cruce entre una taxonomía elemental de los IDS, según los criterios antes citados y una taxonomía de Minería de Datos[26], la cual se poda en la representación a través de sus categorías ortogonales, desplegando como hojas las técnicas, las cuales se correlacionan con las utilizadas en IDS, por medio de la unión de las hojas del árbol.

Por otro lado, se muestra una clase adicional *No Minería de Datos*, para mostrar aquellas técnicas que no entran el espectro de Minería, por el carácter estático de sus planteamientos, los cuales se basan mayoritariamente en la búsqueda de patrones preconcebidos por analistas expertos en seguridad, lo cual es ampliamente usado en productos comerciales.

Clasificación de IDSs revisados A continuación se presentan tablas de clasificación de los IDS bajo estudio, según las taxonomías ya nombradas.

NIDES	Vector de Características
Haystack	Vector de Características
MIDAS	Vector de Características
NSM	Vector de Características
ESMERALD	Vector de Características

Cuadro 1. IDS Basado en Mal Uso y Host, Relacionado con Aprendizaje basado en casos

NADIR	Reglas parametrizadas
ASAX	Reglas
Rosset	C4.5
MADAM ID	RIPPER

Cuadro 2. IDS Basado en Mal Uso y Host, Relacionado con Árboles de Decisión

IDIoT	Redes de Petri Coloreadas
Bro	Sistema Experto Estático
USTAT	Diagrama de Estado de Transición

Cuadro 3. IDS Basado en Mal Uso y Host, Relacionado con Técnica fuera de MD

T. Y. Lin	Conjuntos Difusos
-----------	-------------------

Cuadro 4. IDS Basado en Mal Uso y Red, Relacionado con Agrupaciones Difusas

manHunt	Sistema Experto, anomalía de Protocolo (estático)
Snort	Sistema de Reglas
Nokia IP	Sistema Experto
NID200	Sistema Experto
OneSecure	Anomalía de Protocolo (estático)
StealthWatch M100	Sistema de Reglas
DIDS	Sistema Experto Estático
CISCO IDS 4235	Sistema Experto
SecureNet	Sistema Experto

Cuadro 5. IDS Basado en Mal Uso y Red, Relacionado con Técnica fuera de MD

ESMERALD	Vector de Características
NIDES	Vector de Características
Haystack	Vector de Características
MIDAS	Vector de Características
IDES	Vector de Características
JiNao	Vector de Características
HyperView	Redes Neuronales

Cuadro 6. IDS Basado en Anomalía y Host, Relacionado con Aprendizaje Basado en Casos

MADAM ID	RIPPER
W&S	Sistema Experto con Reglas Autogeneradas

Cuadro 7. IDS Basado en Anomalía y Host, Relacionado con Árboles de Decisión

Forrest	Matching Heurístico
Pinacho	Matching Heurístico

Cuadro 8. IDS Basado en Anomalía y Host, Relacionado con Análisis de Series de Tiempo

ComputerWatch	Solo Monitoreo y Reporte
DPEM	Sistema Experto

Cuadro 9. IDS Basado en Anomalía y Host, Relacionado con Técnica fuera de MD

NSM	Vector de Características
OneSecure	Vector de Características

Cuadro 10. IDS Basado en Anomalía y Red, Relacionado con Aprendizaje Basado en Casos

Bro	Sistema Experto Estático
-----	--------------------------

Cuadro 11. IDS Basado en Anomalía y Red, Relacionado con Técnicas fuera de MD

5. Conclusiones

De la clasificación efectuada se puede concluir que existe una alta cantidad de sistemas que utilizan técnicas no clasificables en el espectro de minería de datos, como lo son los sistemas expertos y otros basados en reglas estáticas o bien en modelos temporales como redes de Petri, los cuales están determinados a priori por un experto analista de seguridad, sin posibilidad de variaciones de forma autónoma. Este es el caso de la mayoría de los sistemas comerciales, los cuales no recurren a técnicas autoadaptivas en sus planteamientos, por otro lado, es de notar, que estos software no son del todo satisfactorios[13]

Por otro lado también existe un gran uso de vectores de características, ya sea en estudios de anomalía o mal uso, especialmente centrados en el host, debido a la gran cantidad de información que de estos se puede rescatar, a partir de sus múltiples logs o bien, del estado en que se encuentran (carga de CPU, número de usuarios, etc.).

Además cabe destacar que la variantes biológicas en el enfoque de detección de intrusos[21][7] presentan usualmente enfoques heurísticos muy apegados al dominio de donde obtienen la información, el que suele ser el host.

Referencias

1. Terran Lane, Carla E. Brodley.: Detecting the Abnormal: Machine Learning in Computer Security (1997),Purdue University.
2. Terran Lane, Carla E Brodley.: An Application of Machine Learning to Anomaly Detection (1997), Purdue University.
3. Terran Lane.: Sequence Matching and Learning in Anomaly Detection for Computer Security (1997), Purdue University.
4. Stuart Russel Peter Norvig, Inteligencia Artificial Un enfoque moderno (1995), Prentice Hall
5. J. S. Balasubramaniya, Jose O. Garcia-Fernandez, David Isacoff, Eugene Spafford, Diego Zamboni.: An Architecture for Intrusion Detection using Autonomous Agents (1998), COAST Laboratory, Purdue University.
6. R. Fielding, C Irvine, J. Gettys, J. Mogul.: RFC 2068 (1997), DEC MIT/LCS
7. Pedro Pinacho D.: Desarrollo de Técnicas para Tratamiento de Intrusión en Sistemas Computacionales, Perspectiva Biológica (2000), Universidad de Concepción.
8. Jeimy Cano.: Credenciales para investigaciones forenses, certificación y entrenamiento, Universidad de los Andes, Colombia, (Agosto 2001).
9. Wenke Lee, A Data Mining Framework for Constructing Features a Models for Intrusion Detection Systems (1999).
10. Sandeep Kumar, A Thesis Submitted to the Faculty of Purdue University, (1995).
11. Stefan Axelsson, Intrusion Detection Systems: A Survey and Taxonomy (2000), Chalmers University of Technology.
12. Aurobindo Sundaram, An Introduction to Intrusion Detection (2000), ACM-CrossRoads.
13. David Newman, Joel Snyder, Rodney Thayer, Crying wolf: False Alarm hide attacks (2002), NetworkWorldFusion.

14. Pedro Pinacho, Ricardo Contreras, Una Propuesta de Sistema para Tratamiento de Intrusos Inspirado en la Biología (2001), Universidad de Santiago, Universidad de Concepción.
15. Pedro Pinacho, Ricardo Contreras, Técnicas de Detección de Intrusos en Sistemas Computacionales Perspectiva Biológica (2000), DIINF-USACH, DIICC-UdeC.
16. Jungwon Kim, An Artificial Immune for Network Intrusion Detection, Department of Computer Science, University College London.
17. Wenke Lee, Salvatore Stolfo, A Data Mining Framework for Adaptive Intrusion Detection (1998), Columbia University.
18. Sandeep Kumar, Eugene Spafford, An Application of Pattern Matching in Intrusion Detection, Purdue University (1994).
19. Tom Fawcett, Foster Provost, Adaptive Fraud Detection (1997), NYNEX.
20. Saharon Rosset, Uzi Murad, Discovery of Fraud for Telecommunications Challenges and Solutions (1999), Amdocs (Israel).
21. Stephanie Forrest, Steven A. Hofmeyr, Computer Immunology (1997), Communications of ACM.
22. , Teresa F. Lunt, Detecting Intruders in Computer Systems, SRI International.
23. , T. Y. Lin, Anomaly Detection, San Jose State University.
24. Steven A. Hofmeyr, Atephanie Forrest, Intrusion Detection using Sequence of System Calls (1997), University of New Mexico.
25. Terran Lane, Carla E. Brodley, Temporal Sequence Learning and Data Reduction for Anomaly Detection, Purdue University.
26. Ricardo Baeza-Yates, Luis F. Bastías, Clasificación y Taxonomía de Minería de Datos (2000).
27. William W Cohen, Learning Trees and Rules with set-valued features.
28. <http://www.cisco.com/warp/public/cc/pd/sqsw/sqidsz/index.shtml>
29. <http://www.snort.org>
30. <http://www.intrusion.com/products/featureandbenefit.asp>
31. <http://www.nokia.com/securitysolutions/network/iss.html>
32. <http://www.onesecure.com/products.html>
33. <http://www.recourse.com/product/ManHunt/>