INSTITUTO POLITÉCNICO NACIONAL

ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA Y ELÉCTRICA

SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

UNIDAD CULHUACAN

"Identificación de ataques de DDoS en redes de datos a través de un modelo basado en una red bayesiana"

T E S I N A

QUE PARA OBTENER EL GRADO DE

ESPECIALISTA EN SEGURIDAD INFORMÁTICA Y TECNOLOGÍAS DE LA INFORMACIÓN

PRESENTA

DR. JOSÉ LUIS TADEO DUARTE ALCÁNTARA

ASESOR:

M. en C. MARCOS ARTURO ROSALES GARCÍA

MEXICO D.F. MAYO, 2011

SIP-14-E

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESINA

En la Ciudad de México, D. F. siendo las 11:30 horas del día 17 del mes de junio del 2011 se reunieron los miembros de la Comisión Revisora de la Tesina, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de SEPI-ESIME-CULHUACAN para examinar la tesina titulada:

"Identificación de Ataques de DDoS en Redes de Datos a través de un Modelo Basado en una Red Bayesiana"

Presentada por el alumno:

Duarte	**Alcántara**	**José Luis Tadeo**
Apellido paterno	Apellido materno	Nombre(s)

Con registro:

B	1	0	1	6	5	1

aspirante de:

ESPECIALIDAD EN SEGURIDAD INFORMÁTICA Y TECNOLOGÍAS DE LA INFORMACIÓN

Después de intercambiar opiniones los miembros de la Comisión manifestaron ***APROBAR LA TESINA***, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Director(a) de tesina

INSTITUTO POLITECNICO NACIONAL ESTADOS UNIDOS MEXICANOS

M. en C. Marcos Arturo Rosales García

Dr. Gualberto Aguilar Torres

S.E.P.
SECCION DE ESTUDIOS DE POSGRADO E INVESTIGACION
ESIME CULHUACAN

Dr. Moisés Salinas Rosales

PRESIDENTE DEL COLEGIO DE PROFESORES

Dr. Gonzalo Isaac Duchen Sánchez

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA CESIÓN DE DERECHOS

En la Ciudad de México, D.F. el día 23 del mes Junio del año 2011, el (la) que suscribe C. José Luis Tadeo Duarte Alcántara alumno (a) del Programa de Especialidad en Seguridad Informática y Tecnologías de Información con número de registro B101651, adscrito a SEPI-ESIME-CULHUACAN, manifiesta que es autor (a) intelectual del presente trabajo de Tesis bajo la dirección de M. en C. Marcos Arturo Rosales García y cede los derechos del trabajo intitulado Identificación de ataques de DDoS en redes de datos a través de un modelo basado en una red bayesiana, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección jduartea68@gmail.com. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Dr. José Luis Tadeo Duarte Alcántara
Nombre y firma

Agradecimientos

Quiero comenzar por agradecerle a Dios que siempre confía en mí, aun cuando muchas veces pierdo la Fe.

Quiero agradecerte Güerita, Rosy mi Esposa por toda tu paciencia, tolerancia, amor, comprensión y apoyo en los momentos tan complicados que hemos enfrentado juntos para poder concluir este proyecto.

A mis dos hermosos hijos, Eric Emmanuel y José Luis Jr., que han hecho sacrificios para regalarme su tiempo para realizar este trabajo.

A mi Madre y a mi Padre, a quienes les debo estar aquí.

Al Maestro Marcos, mi asesor en esta investigación, porque has sido un gran amigo incondicional, principal responsable de que me dieran la oportunidad de ingresar a este posgrado y de que ahora pueda terminar este trabajo, porque creíste en mí desde el principio de esta aventura.

Al Dr. Antonio Castañeda, por tu amistad, tus consejos, por la disposición para regalarme tu basto conocimiento de forma incondicional y porque siempre estás dispuesto a ayudarme, orientarme y enseñarme.

Al Dr. Moisés Salinas, por darme un voto de confianza para hacer este posgrado, por tu amistad y todo el apoyo que me has brindado.

Al Dr. Gabriel Sánchez, por aceptarme en el programa y obsequiarme tu amistad.

A todos los que han sido mis profesores en este posgrado, que me han regalado su conocimiento, paciencia y su tiempo.

A mis compañeros de aulas, todos y cada uno de ellos, me ha regalado su tiempo, apoyo, solidaridad y conocimientos, de ellos he aprendido el valor de la amistad.

Y por supuesto gracias a esta grande y noble institución y a sus autoridades.

TABLA DE CONTENIDO

Página

LISTA DE FIGURAS

Página

RESUMEN

Este trabajo presenta el diseño y evaluación en un ambiente experimental, de un modelo basado en redes bayesianas, para la identificación de un ataque por denegación de servicio distribuido (DDoS) a una infraestructura de una red de datos.

ABSTRACT

This research presents the design and evaluation in an experimental environment, of a model based upon bayesian networks for the identification of an attack by distributed denial of service (DDoS) to a network infraestructure.

PALABRAS CLAVES:

Redes Bayesianas, Ataque DDoS, Redes de Datos.

KEYWORDS:

Bayesian Networks, DDoS Attack, Data Networks.

1. INTRODUCCIÓN

Este trabajo propone un modelo para la identificación de ataques de DDoS (por sus siglas en inglés, Distributed Denial of Services) en redes de datos, a través de la construcción de una red bayesiana que permite identificar variables aleatorias y sus respectivos estados para calcular posteriormente su inferencia y lograr entonces un diagnóstico de anomalías y ataques en la red de una organización.

La investigación ha iniciado con una revisión del marco teórico de las redes bayesianas, revisando brevemente su origen, características en su estructura y formación, los principios que las rigen, los tipos de topologías y probabilidades con las que pueden formarse y los modelos de inferencia para hacer que propongan una estimación y pronóstico de algunos elementos de interés.

Posteriormente se ofrece una revisión del estado del arte del análisis forense para redes de datos que sufren ataques de este tipo, así como la definición de las necesidades de identificación, clasificación y recolección de evidencia que permita, en una visión amplia, una análisis forense completo y suficiente sobre este tipo de sucesos en las redes de datos de las organizaciones.

En seguida se propone un diseño de investigación, se describen las hipótesis correspondientes así como la justificación, viabilidad y aportaciones del presente trabajo.

Inmediatamente se presenta el diseño del modelo de la red bayesiana para la identificación de ataques de DDoS, se explica el esquema de evaluación del modelo propuesto y se exponen los resultados obtenidos a través de una red experimental que se utilizó en esta investigación.

Se exponen también las recomendaciones que se derivaron de la experiencia obtenida a lo largo del diseño, desarrollo y realización de la investigación así como los retos y limitaciones que se enfrentaron a lo largo de la misma.

Finalmente se presentan las conclusiones a las que se llegaron en el presente trabajo, basadas completamente en los resultados obtenidos previamente, con lo que

podemos confrontar y comprobar todas y cada una de las hipótesis de esta investigación así como la construcción de nuevo conocimiento producto de este trabajo.

1.1. Antecedentes

Los ataques a redes de datos a través de gusanos (Worms) y por denegación de servicio distribuido (DDoS) han sido en los últimos 10 años las amenazas principales a las redes de datos e Internet.

Debido al desarrollo de hardware, los servidores web de algunos departamentos, áreas o funciones claves para las organizaciones, como puede ser el comercio electrónico, redes de datos en instituciones financieras y gubernamentales, han mostrado una tolerancia amplia respecto al ataque de gusanos, pero no así al ataque por DDoS.

Muchos modelos de identificación han sido propuestos desde el año 2000, cuando se conoció el primer caso de un ataque por DDoS, pero siguen desarrollándose investigación para encontrar nuevos modelos para aproximarse a una identificación amplia de este ataque.

Aun cuando empresas de tecnologías de información y telecomunicaciones han desarrollado una serie de productos y servicios para ofrecer seguridad a las organizaciones, no se ha logrado detener o reducir significativamente este tipo de ataques.

A penas en el año 2010, el gobierno español logró identificar algunos presuntos responsables del ataque más grande, por la cantidad de equipos utilizados para realizarlo, registrado en los años recientes por DDoS, conocido como la "*botnet Mariposa*".

La mayor parte de la incapacidad de enfrentar a la justicia a los responsables de este tipo de ataques, es por la falta de evidencia completa y suficiente en las redes de datos y servidores víctimas, debido a la deficiente identificación de un ataque de DDoS.

Es claro que los recursos económicos para construir una red de datos que grabe todo el tiempo el tráfico que existe en su infraestructura, son escasos e insuficientes en las mayorías de las organizaciones, y también es claro que aun cuando se lograra recopilar

esta cantidad de información monumental, no se tendría la capacidad de analizarla para identificar un ataque a tiempo.

Justamente en este paradigma se encuentran la mayoría de las organizaciones con una preocupación en la seguridad de sus activos tecnológicos y con la imperiosa necesidad de encontrar soluciones viables en la parte económica, operativa y tecnológica.

1.2. Planteamiento del problema

La amplia y cada vez más continua utilización de las redes de datos para casi cualquier actividad humana, ha originado que cada vez más personas, organizaciones, corporaciones, naciones y gobiernos dependan de la infraestructura de comunicaciones y sus interconexiones para realizar sus actividades regulares y de misión crítica.

Esto ha llevado a todos los usuarios de estas infraestructuras de comunicaciones, a la necesidad de proteger sus activos informáticos, que en muchas ocasiones, también son estratégicos para sus propósitos y objetivos.

De forma paralela también se ha desarrollado ampliamente el interés de individuos, grupos económicos, políticos, sociales e incluso naciones de utilizar, apoderarse, sabotear y atacar estos activos informáticos, en la búsqueda de sus propios fines y objetivos perversos.

En el 2009 el US-CERT (por sus siglas en inglés, United States- Computer Emergency Readiness Team) recibió reportes de los usuarios y administradores de servicios de redes en Estados Unidos, sobre incidencias, anomalías y ataques a sus activos informáticos; Donde el 73.4% fueron búsquedas de acceso a recursos (scans), identificación de puertos, protocolos, servicios o cualquier combinación para la explotación de vulnerabilidades (Probes or Attempted Access) en sus equipos.

La estadística que presenta el US-CERT en el 2009 se basa en notificaciones que realizan las entidades federales de los Estados Unidos, Corporaciones y Empresas, así

como individuos o pequeños grupos de interés que de alguna forma han logrado detectar este tipo de incidencias o ataques.

Lo que nos lleva a considerar que el porcentaje de ataques puede ser sensiblemente mayor a lo reportado, considerando que no todas las organizaciones, empresas, grupos o individuos cuentan con herramientas que les permitan detectar este tipo de incidencias en sus activos informáticos.

La ICCC (por sus siglas en inglés, Internet Crime Complaint Center) que depende del FBI (por sus siglas en inglés, Federal Bureau of Investigation) y del Department of Homeland Security, han informado que en el 2010 han recibido 2 millones de denuncias de ciber-crimen (cyber crime), lo que da un promedio de 25,000 incidencias criminales en internet cada mes.

Esta organización dentro de sus funciones está la de analizar, verificar, compilar la información respectiva para poder llevar a una corte federal de los Estados Unidos este tipo de denuncias. En el 2010 informó que sus analistas lograron preparar 1,420 casos, los cuales correspondieron a aproximadamente 49,808 denuncias. El FBI logró presentar solamente 698 casos en cortes federales de los 1,420 que le fueron entregados. Esto debido a la falta de evidencia sustancial de los ataques o incidentes.

Los ataques perpetrados a redes de datos, presentan este tipo de eventualidades y desventajas, su recolección de evidencia es pobre, altamente volátil y en varias ocasiones no es posible el rastreo por las técnicas de suplantación de credenciales digitales.

Aun cuando existen equipos y software de monitoreo de redes, que pudieran permitir registrar todas las actividades que se realizan en la infraestructura de comunicaciones, las organizaciones enfrentan los siguientes problemas:

- Cuando la red tiene una importante cantidad de nodos e interconexiones, el presupuesto requerido para la compra, implantación y monitoreo de toda la red resulta inviable para casi cualquier organización.

- Aun cuando se pudiera monitorear una gran cantidad de puntos en la red de datos, el tiempo de análisis de todas y cada una de las bitácoras de los equipos

y software sensores de la red, hacen prácticamente imposible identificar alguna incidencia o posible ataque.

- De cualquier forma, estos equipos y software de monitoreo de redes, consumen una gran cantidad de recursos de la red, que pueden llegar a originar un desempeño pobre de la misma.

- En el caso de ser selectivo en la ubicación de los sensores y monitores del comportamiento de la red, se requiere del conocimiento de expertos en el comportamiento de cada una de las redes monitoreadas, que ayude a la determinación e identificación de alguna anomalía o incidencia, lo que hace también inviable la posibilidad de identificar una incidencia en el momento en que sucede.

Por lo anterior, regularmente los análisis forenses, que de forma estructural se hacen *post mortem*, pueden dar como resultado la no identificación de incidentes o ataques, puesto que una buena parte de la información se pierde una vez que el ataque deja de suceder, y solamente se cuenta con lo que se haya logrado recolectar si y solo si, un equipo de monitoreo fue colocado en un punto tal que haya permitido registrar las actividades relativas al ataque.

Esto nos lleva a considerar que un modelo ligero como el que ofrecen las redes bayesianas pueden ayudar a la identificación en tiempo de real de alguna anomalía o incidencia en la red de datos en tiempo real, sin incurrir en las problemáticas que arriba se describieron.

El modelo generado a través de una red bayesiana, puede ser instalado en múltiples puntos en la topología de la red, para efecto de poder identificar diferentes tráficos y sus comportamientos anómalos de forma combinada, que si bien, por ser una red bayesiana múltiple o un poli árbol, su procesamiento inferencial puede tomar más tiempo

Aun cuando el análisis inferencial completo de la red bayesiana consuma un tiempo mayor, se puede definir un umbral, para que la red de la instrucción a algún dispositivo o incluso a una HoneyNet para iniciar la recolección de evidencia.

En el caso de confirmar un falso positivo después del análisis, de igual forma se puede dar la instrucción al dispositivo que suspenda la recolección de evidencia y así optimizar los recursos informáticos.

1.3. Objetivos

Este trabajo tiene como objetivo, proponer los primeros elementos para el desarrollo de una línea de investigación, que identifique incidencias y anomalías en el tráfico de una red, a través de la construcción de un modelo de inferencia basado en la teoría de las redes bayesianas.

1.3.1. Objetivo general

El objetivo general de este trabajo es lograr la identificación de incidencias en el tráfico de una red de datos a través de una red bayesiana.

1.3.2. Objetivos específicos

Desarrollar un modelo basado en redes bayesianas, que permita identificar incidencias, anomalías y un posible ataque de DDoS en tiempo real.

Identificar, medir y calibrar las variables aleatorias de una red bayesiana, que se involucran durante un ataque de DDoS a una red de datos.

Evaluar y ajustar los parámetros que caracterizan a una red bayesiana una vez que se identificó alguna anomalía o incidencia en el tráfico de una red de datos.

1.4. Justificación

La justificación de esta investigación proviene de varios elementos que conjugados, le dan sentido al objetivo de este trabajo:

No hay una amplia cantidad de investigaciones sobre la detección de anomalías en redes de datos utilizando redes bayesianas, que sin embargo, son utilizadas como modelos de predicción en distintos casos de en la ciencia forense. (Taroni, Bozza & Aitken, 2005; Taroni, Bozza & Biedermann, 2005; Aitken, Gammerman, Zhang, Connolly, Bailey et. Tal, 1996)

La constante tasa de crecimiento en el uso de redes de datos de forma, personal, empresarial, corporativa y multinacional, ofrece una infraestructura cada vez más grande para realizar ataques a objetivos al azar o predeterminados por cualquier individuo, empresa, corporación, nación o bloque de naciones. (ICCC, 2010)

No existe una gran cantidad de investigaciones que propongan la utilización de un modelo inferencial de probabilidad a través de una técnica de grafos que pueda analizar en tiempo real, flujos de paquetes de datos para identificar alguna incidencia, anomalía o posible ataque, que dé inicio al proceso de recolección de evidencia que sirva al análisis forense.

2. MARCO TEÓRICO

2.1. Estado del arte del análisis forense en red

Una serie de componentes y equipos de seguridad de red, como son los Firewalls y los IDD's, por sus siglas en inglés (Intrusion Detection Devices) son el primer frente activo contra una gran cantidad de ataques a redes en todas partes del mundo, de parte de virus, gusanos, scanners y ataques más elaborados en cuanto a su proceso de ejecución como son el ataque por denegación de servicio (DDoS).

Actualmente se tienen modelos y métodos para identificación de firmas digitales de paquetes en particular los cuales, permiten identificar patrones de secuencia de ataques conocidos así como la posibilidad de identificar origen y ubicación física del posible ataque.

Sin embargo, regularmente los atacantes expertos, logran cambiar sus credenciales de identificación y localización para evitar ser ubicados, por lo que regularmente lo que se puede obtener es solo el modelo de ataque.

Por supuesto que la identificación a través de firmas está limitada al trabajo de identificación, análisis y desarrollo de las empresas que ofrecen software y hardware de seguridad, por lo que, anomalías o firmas digitales nuevas o desconocidas, no podrían ser detectadas ni atendidas por estos componentes.

Se puede considerar como un anomalía, a un tipo de ataque conocido que haya podido ser modificado de alguna forma para evitar su detección o también podría presentar de la forma de un ataque completamente nuevo, concebido desde cero.

Los esfuerzos sobre la detección de anomalías en la información que fluye en las redes de datos de las organizaciones, ha llevado a una serie de desarrollos de investigación utilizando métodos provenientes del análisis estadístico de señales, teoría de reconocimiento de patrones y modelos predictivos a través de series de tiempo, entre varias otras técnicas y métodos.

En este entorno, Yurcik & Le (2005) y Plonka (2000), han demostrado que el concepto conocido como Flujo de Red (Net Flow) y Escaneo o Exploración del Flujo (Flow Scans) puede ser una forma altamente eficiente para monitoreo de una red y análisis de su comportamiento. Gong (2004A; 2004B), propone una metodología en la que el Flujo de Red puede ser utilizado para detectar gusanos y otro tipo de intrusión en la red.

Sang & Li (2000) explican como la tendencia en los fabricantes de componentes de seguridad, es hacia la predicción del tráfico en una red empleando el modelo ARMA (ARMA, por sus siglas en inglés, AutoRegressive Moving Average) también conocido como el modelo de Box-Jenkings, por el tipo de metodología que se utiliza para la estimación del procesamiento de señales, que regularmente se aplica en datos de series de tiempo correlacionados, como sería el caso del tráfico en la red con respecto al tiempo.

De forma paralela Cho, Kaizaki & Kato (2002) exponen en su trabajo un método en el cual el tráfico de datos cercano a en línea, puede ser medido y filtrado utilizando el algoritmo llamado *Patricia Tree* y su extensión conocida como ISGF (por sus siglas en inglés, Inverse Stack Growth Function).

En el trabajo de Pang, Yegneswaran, Barford, Paxson & Peterson (2004), exponen anomalías conocidas y algunos posibles líneas de detección a través del filtrado de datos para reducir los costos de carga en la infraestructura de tecnológica.

Kwitt & Hoffman (2007) así como Shen, Chen & Qin (2007) presentan trabajos relacionados con detección de anomalías utilizando PCA (por sus siglas en inglés, Principal Component Analysis) y mediciones sobre el entorno donde opera la red. Este análisis de anomalías se complementa con la detección de *botnets* a través del trabajo de Karasaridis, Rexroad & Hoeflin (2007).

Existen otros trabajos sobre la detección de intrusiones a redes utilizando la creación de espacios de covarianza basado en métodos de reconocimiento de patrones como pueden ser las investigaciones de Jin, Yeung & Wang (2007).

Feldman, Gilbert & Willinger (1998) a su vez, propone un enfoque basado en una búsqueda en cascada que atiende el tráfico de una red como un comportamiento multifractal, es decir, como un gran conjunto de datos de tráfico formado a su vez por subconjuntos jerárquicos de diferentes topologías (fractales) que pueden ser obtenidas a través de métricas específicas.

Existe otra corriente de investigaciones que ha desarrollado una medición para cuantificar el tráfico de información en redes y han sido ampliamente estudiados para detección y prevención de anomalías, conocida como *grado de entropía* o simplemente *entropía*, tanto Gu, McCallum & Towsley (2005) como Lakhina, Crovella & Diot (2005) exponen desarrollos para obtención de la *entropía* basado en un análisis del tráfico en redes.

Existen investigaciones sobre el estudio de la estructura y flujo de tráfico en redes en conjunto con la correlación visual de las alertas que se presentan en una red. (Foresti, Agutter, Livnat, Moon & Erbacher, 2006).

Otras investigaciones tiene el enfoque del análisis de la entropía del tráfico de una red a través de una revisión estadística longitudinal en el espacio del tiempo. (Eimann, Speidel, Brownlee & Yang, 2005; Lall, Sekar, Ogihata, Xu & Zhang, 2005; Harrington, 2006; Gianvecchio & Wang, 2007).

En el trabajo de Harrington (2006), se desarrolla el manejo de la entropía como una distribución de segundo orden para detectar cambios en el comportamiento del tráfico de la red y Lall, Sekar, Ogihata, Xu & Zhang (2005), utiliza la entropía de la distribución del tráfico como ayuda en el monitoreo de las redes. Ahora bien, Gu, McCallum & Towsley (2005), utilizan la entropía para medir y detectar cualquier tipo de anomalía que puede presentar las redes de alta disponibilidad.

En el trabajo de Kim, Reddy & Vannucci (2004) los encabezados de los paquetes de datos son examinados utilizando análisis agregado de correlación de datos y transformación de señales discretas. La utilización del análisis estadística de datos así como la teoría de reconocimiento de patrones son aplicados para resolver el mismo

problema de identificación de anomalías en el comportamiento del tráfico de la red variando los grados de éxito en la detección de ataques.

Wagner & Plattner (2005), presentan en su trabajo una discusión de un método basad en los cambios en el contenido de la entropía por las direcciones IP's y los puertos identificados, pero no han intentado identificar el tráfico normal del que puede considerarse extraño o subnormal.

El trabajo que presenta Thottan & Ji (2003), aplican técnicas de procesamiento de señales para detectar anormalidades en la red, utilizando análisis estadístico de datos, particularmente identifica la IP de la red que presenta alguna anormalidad definiendo dominios de clase simple en unión con los tipos de fuentes de datos disponibles para realizar el análisis.

Hajji (2005) fundamente su trabajo abordando el problema de la identificación de los cambios en las características en el tráfico de la red, dando mayor énfasis en la detección rápida para reducir los impactos potenciales de los problemas en los servicios que la red provee, esto lo logra utilizando el modelo Gaussiano de tráfico mixto, GMTM (por sus siglas en inglés, Gaussian Mixture Traffic Model).

Uno de los objetivos elementales para realizar un análisis forense es la capacidad de reducción de falsos positivos en la identificación de anomalías de la red, se ha logrado parte de este objetivo, a través de la medición de la entropía de Tsallis, que supera la técnica de medición de entropía de Shannon. (Ziviani, Monsores, Rodrigues & Gomes, 2007)

La técnica de análisis de series de tiempo también es utilizada para la detección de anomalías. Esta técnica de análisis en diferentes encabezados de paquetes de datos y los mecanismos de eficiencia para colectar y analizar datos en tiempo real es uno de los resultados descritos en el trabajo de Kim & Reddy (2008). También demuestran que la serie de tiempo propuesta tiene alta eficacia en la detección de ataques respecto a la técnica de análisis del volumen del tráfico en la red.

El trabajo de Kang & Kim (2008) propone la identificación del tráfico en red malicioso a través de un modelo que genera una visualización utilizando gráficos con códigos de colores, que describan el protocolo recibido en los paquetes así como dos conjuntos de datos adicionales todo clasificados en una línea de tiempo, de tal forma, que se pueda observar la densidad gráfica del comportamiento del tráfico de la red y pueda entonces identificarse manualmente una posible anomalía.

La investigación de Nehinbe (2010), expone la necesidad de tener dispositivos de detección de intrusos en la red NIDS (por sus siglas en inglés, Network Intrusion Detection System) para poder entonces realizar un análisis forense y una auditoría a la red a través de la revisión de la bitácora de incidentes que éstos dispositivos poseen. Este trabajo propone la adición de filtros por subconjunto de datos después de los dispositivos físicos de detección a través del uso del IDS (por sus siglas en inglés, Intrusion Detection Sytems) Snort. Esto permitiría realizar clustering de diferentes criterios sobre el análisis de la bitácora respectiva de forma off-line.

El trabajo que presentan El-Shehaly, Gracanin, Abdel-Hamid & Matkovic (2009), hacen una propuesta de un sistema de conocimiento internet-level para ejecutar comandos de análisis de rastreo en redes, que puede ofrecer una interface GUI para la visualización del resultado del rastreo, identificación de patrones de interés en el tráfico de la red y actividades relacionadas a la identificación de evidencia.

Rekhis, Krichene & Boudriga (2009), realizaron un trabajo donde proponen un sistemas que le llaman DigForNet, el cual tiene como propósito analizar incidentes de seguridad y la explicación de los pasos que utilizaron los atacantes, sin embargo, su base de conocimiento esta determinado por un equipo de expertos en seguridad y ataques, que deben analizar a través una serie de herramientas, un conjunto de escenarios de posibles ataques que han sido recolectados a través de las bitácoras de equipos que se encuentran instalados en algunos puntos de la red.

Este trabajo no alcanza a explicar algunas consideraciones como: cuales son las condiciones o premisas para la ubicación física de los equipos que registrarán los incidentes en la red, cual es el margen de tiempo para revisar estas bitácoras por el equipo

de expertos en seguridad y ataques, cual es el tiempo que se toma el equipo de expertos en analizar la información recopilada en la red y una vez que determinan que sucedió un ataque, cómo reconstruyen la evidencia del mismo, puesto que existe la posibilidad de que el atacante ya no esté ejecutando ningún evento y quizá también haya borrado sus huellas.

En el trabajo que realizó Lin, Zhitang & Cuixia (2009), se propone la recolección de evidencia de algún ataque a través de la ubicación física de una serie de sensores a lo largo y ancho de la red que se desea asegurar, así mismo propone una serie de algoritmos para ofrecer una mayor calidad en la información que se recolectó a través de estos sensores. De igual forma ofrece una reconstrucción del escenario original del ataque a través de una metodología propuesta por los autores, y un algoritmo de agregación de nuevos patrones identificados en una base de conocimiento que pueda incrementar el nivel de automatización en la identificación de un ataque.

Ming & LiZhong (2009) por su parte, presentan un modelo para la detección de una invasión en la red y al mismo tiempo recolectar información del dispositivo atacado en la red. Esta investigación, solo propone un modulo de colección de datos provenientes de las estructuras de paquetes TCP, UDP e IP que posteriormente es comparada con una base de datos donde se tienen patrones registrados de ataques conocidos que se compararán con los obtenidos en la recolección y en caso de no encontrar ninguna correlación, se procederá a determinar si es un ataque o no, por parte de un experto de forma off-line, para entonces decidir si se ingresa esta nuevo patrón a la base de conocimiento o solamente fue un falso positivo.

Amran, Phan & Parish (2009) proponen crear un modelo intuitivo de recolección de evidencia de ataques en una red, a través del análisis simple de paquetes de datos que pueden ser considerados sospechosos, asignando una calificación a cada paquete que es considerado candidato a ser analizado. Esta calificación que se asigna esta en función de la severidad del tipo de ataque que se cometió y ofrece una medida del grado de credibilidad que la evidencia recolectada puede tener.

El modelo intuitivo que se propone en este trabajo, está relacionado con la recolección de evidencia que es analizada a través de tres indicadores simples como son:

ataque pasivo, ataque activo y ataque intrusivo. Una vez que se califica este paquete entonces se buscará completar la información para tener una evidencia amplia, como puede ser, qué se atacó, cuando fue atacado, desde dónde se realizó el ataque, etc.

Este esfuerzo que presenta esta investigación, esta bajo el supuesto de que se está recolectando todo el tiempo paquetes de la red que son sometidos a el análisis simple que ya fue comentado, para luego realizar la búsqueda de evidencia *post mortem*, lo cual solo permite analizar eventos que ya sucedieron y con ello prevenir el futuro ataque.

Algo similar ocurre con el trabajo de Achi, Hellany & Nagrial (2008), quienes proponen esquemas de recolección de evidencia para el futuro trabajo de análisis forense a través, de la instalación de una serie de equipos de seguridad de proveedores conocidos, que van desde dispositivos Firewalls hasta ASA's (por sus siglas en inglés, Adptative Security Appliances), combinado con esquemas de diseño de redes seguras proporcionadas por los fabricantes más importantes en la industria de las telecomunicaciones.

Ninguno de los estudios anteriores proporciona un modelo que permita identificar una posible amenaza en tiempo real para entonces comenzar la recolección de información relativa al posible ataque y entonces el atacante pueda ser redirigido a una zona de monitoreo integral en la red, como puede ser una HoneyNet, para continuar recolectando información on-line del comportamiento del atacante, sus técnicas, herramientas, ubicación lógica, etc.

2.2. Redes Bayesianas

Una red bayesiana es una forma de representación de gráfica sobre las relaciones que pueden observarse entre variables aleatorias y la dependencia directa de éstas.

Las variables aleatorias entonces, son representadas en forma gráfica, como nodos y las líneas que conectan estos nodos, representarán las dependencias existentes entre los mismos.

Se entiende como topología de la red bayesiana, a la estructura que toma la red, la cual nos ofrece información de la probabilidad en cada una de las dependencias que existe con respecto a las variables aleatorias.

Estas redes, de igual forma, pueden representar las independencias condicionadas de una o varias variables aleatorias una vez que estén dadas una o varias variables aleatorias de manera previa. Esto es, una relación condicionada independiente se presenta cuando una variable no está relacionada directamente con otra variable, esto puede extenderse considerando que la red esté formada por un número grande de variables aleatorias.

La Figura 2.3.1 muestra el caso de relaciones condicionales independientes, donde las variables D, E y F guardan una relación condicional independiente con la variable C.

Podemos expresar estas relaciones con una notación de probabilidad condicional, quedando como sigue:

$$P\left(C\middle|A,B,D,E,F\right)=P\left(C\middle|B\right) \tag{2.3.1}$$

Lo que nos muestra que la variable aleatoria B esta separando a la variable C del resto de la estructura de la red bayesiana.

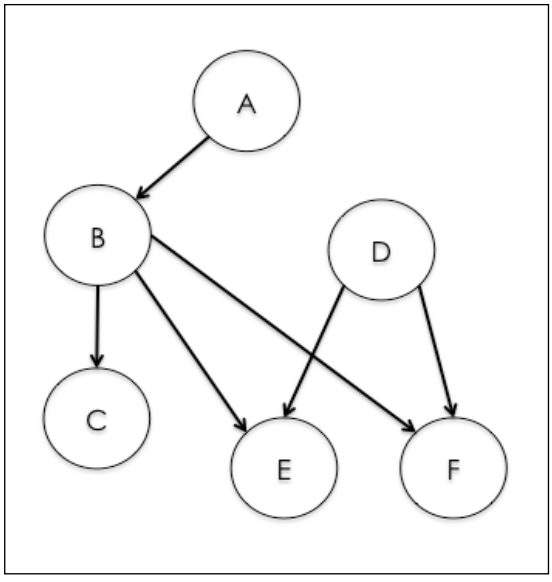

Figura 2.3.1 Representación gráfica de una red bayesiana.

Identificar las relaciones de independencia en las redes bayesianas representa, la posibilidad de simplificar su topología gráfica para representar conocimiento, obteniendo entonces, una estructura con menos parámetros y una propagación más sencilla de las correspondientes probabilidades.

Las redes bayesianas permiten mostrar las dependencias e independencias entre variables aleatorias, calificando la dispersión de probabilidad a partir de las independencias condicionadas de acuerdo al teorema de Bayes.

Una notación para exponer las independencias condicionales de X que es independiente de Y dado la variable Z, es la que a continuación se presenta:

$$P(X|Y,Z) = P(X|Z) \tag{2.3.2}$$

Se tienen al menos tres tipos de grafos que permiten describir algunas relaciones entre los nodos en una red bayesiana:

- Nodos de seriales: A $\rightarrow$B $\rightarrow$C.
- Nodos divergentes: A $\leftarrow$B $\rightarrow$C.
- Nodos convergentes: A $\rightarrow$B $\leftarrow$C.

Cuando se tiene evidencia con un alto grado de certeza en alguno de los estados que guarda una red bayesiana (nodo), se dice que la variable aleatoria se encuentra *instanciada,* esto permite realizar una serie de análisis a una red causal, que reduce el tamaño de la misma, permitiendo entonces, hacer algoritmos de propagación de probabilidades más sencillos y rápidos.

Para la evaluación de la independencia condicional se utiliza el criterio llamado *separación-D* (*d-separation*), el cual se define como sigue:

Se puede decir que el conjunto de variables *A* es independiente del conjunto de variables *B* dado el conjunto *C*, en una red causal, si no existe ninguna trayectoria entre *A* y *B* en que:

1. Todos los nodos convergentes están o tienen descendientes en *C*.

2. Todos los demás nodos no están en *C*.

Las condiciones anteriores, Jensen & Nielsen (2007) las expresa en términos de la *instanciación* de las variables aleatorias, exponiendo que la conexión entre *A* y *B* debe ser serial y que la variable *C* se encuentra *instanciada*. Así mismo, define que la conexión deberá ser convergente y ni la variable *C* ni ninguno de sus descendientes tiene ningún tipo de evidencia recibida, es decir, no estarán *instanciadas*.

En la Figura 2.3.2 podemos observar que las variables *B* y *M* se encuentran instanciadas, por lo que podemos asegurar bajo la definición de la *separación-D*, la variable *A* se tiene una *separación-D* exclusivamente de la variable G.

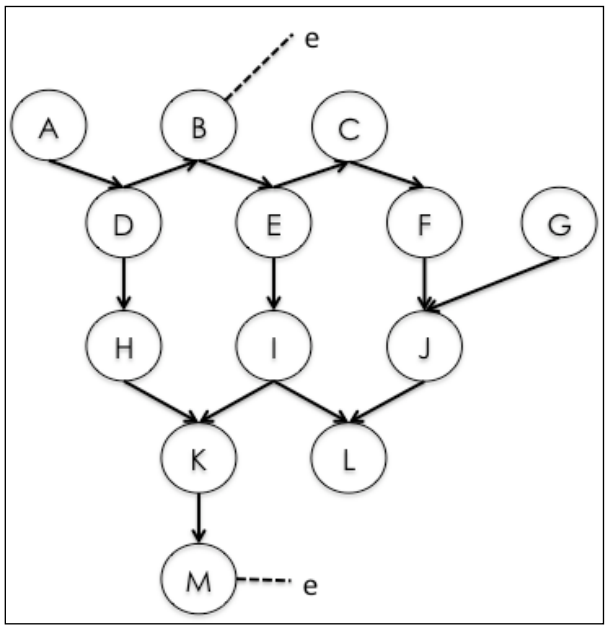

Figura 2.3.2 Representación gráfica de una red bayesiana con B y M instanciadas. (Jensen & Nielsen,2007)

Para exponer cuales son las variables que han recibido evidencia certera, se denota con un arco en línea punteada con la terminación de la letra *e* de evidencia.

Cuando dos variables *A* y *B*, no tienen *separación-D*, se considera que se encuentran *conectadas-D*, sin embargo es importante resaltar, que aun cuando esta condición se presente, no necesariamente los cambios que experimente la variable *A* impactarán en la variable *B*.

En este momento es importante presentar un elemento útil para la construcción y análisis de redes causales, conocido como *Markov Blanquet*, el cual definiremos como sigue:

La *Markov Blanquet* de una variable *A* es el conjunto formado por los padres de *A*, los hijos de *A* y las variables que comparten un hijo con *A*.

En el caso en que todas las variables de una *Markov Blanquet* de *A* estuvieran *instanciados*, entonces, *A* estará *separada-D* del resto de la red causal.

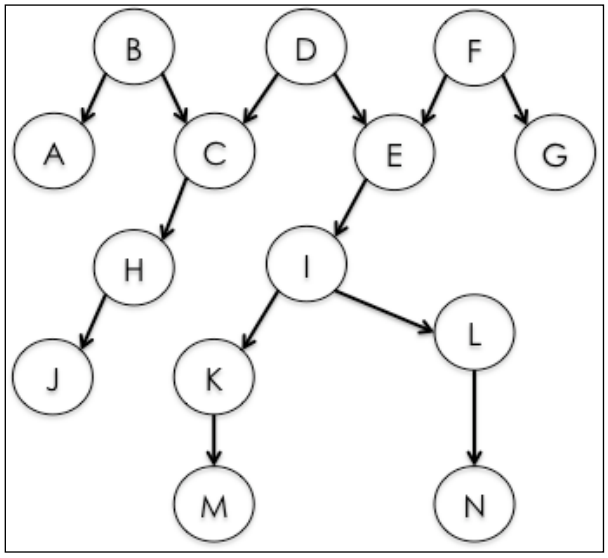

Figura 2.3.3 Representación gráfica de una Markov Blanquet. (Jensen & Nielsen,2007)

Podemos observar en la Figura 2.3.3. que la *Markov Blanquet* para la variables aleatoria *I* es $\{C,E,H,K,L\}$, en el supuesto que todas las variables "vecinas" de *I*, estuvieran *instanciadas*, entonces, podríamos afirmar que *J* no está *separada-D* de *I*.

Se puede decir entonces, que una red bayesiana es un grafo a cíclico dirigido, el cual describe la distribución de probabilidad conjunta que existe en un conjunto de variables aleatorias.

Consideremos que P = $\{X_1,X_2,X_3...,X_n\}$ son un conjunto de variables aleatorias, una red bayesiana *RB* definirá una distribución de probabilidad única sobre *P* dada por la siguiente ecuación:

$$P_{RB} = \left(X_1, X_2, X_3, ..., X_n\right) = \prod_{i=1}^{n} P_{RB}\left(X_i \middle| \prod X_i\right) \tag{2.3.3}$$

Podemos también plantear la regla fundamental del cálculo de probabilidades:

$$P\left(A|B,C\right)P\left(B|C\right) = P\left(A,B|C\right) \tag{2.3.4}$$

Las redes bayesianas parten de la regla de Bayes, la cual se define como:

$$P\left(B|A,C = \frac{P\left(A|B,C\right)P\left(B|C\right)}{P\left(A|C\right)}\right) \tag{2.3.5}$$

La condicional de independencia que se utiliza para las redes bayesianas se define como:

A y C son independientes dado B si $P\left(A|B\right) = P\left(A|B,C\right)$ (2.3.6)

La separación-D permiten para las redes bayesianas plantear que si *A* y *B* tienen una *separación-D* en una red bayesiana y se cuenta con evidencia *e* definida, entonces:

$$P\left(A|B,e\right) = P\left(A|e\right) \tag{2.3.7}$$

El considerar las relaciones de independencia en la topología de las redes bayesianas, permite que sean una herramienta sencilla para representar conocimiento por que se puede realizar de forma compacta, por la reducción de parámetros para expresar un grupo de conocimiento. Adicionalmente, proporcionan métodos flexibles basados en la propagación de la distribución de probabilidades por toda la red, basado principalmente, en las ecuaciones 2.3.4, 2.3.5, 2.3.6 y por supuesto la 2.3.7.

Existen una serie de técnicas para lograr que el tamaño del número de parámetros y sus correspondientes tablas de probabilidad no crezca de forma exponencial por el número de padres que cada nodo o variables aleatoria puede tener en una red.

El objetivo es reducir lo más posible, el número de parámetros para concentrarnos en exclusivamente los necesarios para poder operar y analizar una red bayesiana. Regularmente se utilizan para este fin, los modelos canónicos como:

- Noisy OR (interacción disyuntiva)
- Noisy AND (interacción conjuntiva)
- Noisy Max Gate
- Noisy Min Gate

Cada modelo se diseña para un caso en particular respecto al tipo de relaciones que guarda una red. Sin embargo, regularmente se utiliza el modelo *Noisy OR*, el cual es aplicable cuando se tiene que varias causas pueden derivar un efecto cada una por si sola, pero la probabilidad del efecto no se ve afectada negativamente si se presentan de forma concurrente más de una causa.

Lo que significa que cuando se presenta el uso del Noisy OR, solo se especificará un parámetro por cada nodo padre, en vez de 2^n siendo n el número de padres y considerando un árbol con topología binaria.

Cuando ya se tiene una red y una serie de valores específicos de algunas de las variables aleatorias que lo conforman, es posible comenzar a realizar estimaciones de los valores del resto de las variables de la red, aplicando un razonamiento probabilístico.

Esto es, sobre las redes bayesianas, se propagan los efectos de las evidencias concretas (variables aleatorias conocidas) para conocer las probabilidades a *posteriori* de las variables desconocidas. De forma general, una red puede utilizarse para calcular la distribución de probabilidad para cualquier conjunto o subconjunto de variables aleatorias, a partir de los valores que asumen un conjunto o subconjunto de variables restantes.

En la Tabla 2.3.1 se presenta una breve clasificación de los tipos de algoritmos que se pueden utilizar para efecto de hacer estas inferencias en una red bayesiana, como es de esperarse, la elección del algoritmo estará en función de la topología de la red y el número de variables objetivo que se desee encontrar su probabilidad.

El proceso de inferencia para el caso de las redes bayesianas tiene como objetivo, determinar la estimación (inferencia) correspondiente de aquellas variables desconocidas en la red, a las que llamaremos *clase*, por lo que entonces podemos afirmar que, la inferencia permite generar clasificadores para la red.

Tabla 2.3.1 Clasificación de tipos de algoritmos de propagación para redes Bayesianas

Topología	**No. Variables**	**Algoritmo**
Cualquier grafo	Una variable	Eliminación
Estructuras de conexión sencilla, árboles o poli árboles	Cualquier número de variables	Propagación de Perl
		Agrupamiento (Junction Tree)
Cualquier grafo	Cualquier número de variables	Simulación Estocástica
		Condicionamiento

Fuente: (Jensen & Nielsen, 2007; Taroni, Aitken, Garbolino & Biedermann, 2010) Adaptado por José Luis T. Duarte Alcántara

La inferencia inicia considerando que en la red bayesiana se conocen todos los atributos, para entonces, realizar la clasificación infiriendo (suponiendo) la probabilidad posterior de los valores de cada una de las clases (variables desconocidas) definidas en la red y eligiendo el valor que logre la maximización de la probabilidad.

Un clasificador que es altamente utilizado por su sencillez es el Naive Bayes, el cual tiene como suposición fundamental que todos los atributos A_n (nodos de la red) son independientes entre sí conocida la clase C, por lo que no existen arcos entre los atributos y existe un arco nodo raíz a cada uno de los atributos. La Figura 2.3.4 muestra una red del clasificador Naive Bayes.

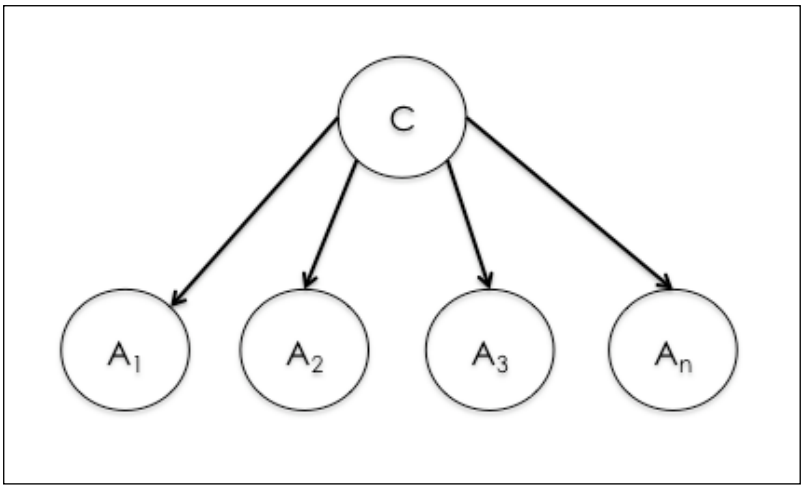

Figura 2.3.4 Representación gráfica del clasificador Naive Bayes

En la Figura 2.3.5 se muestra un ejemplo sencillo de una red bayesiana que representa un grado de conocimiento sobre dos tipos de ataques informáticos, como son el de *DoS* (por sus siglas en inglés, Denial of Services) y el de *ARP Poisoning* (envenenamiento de paquetes ARP, por sus siglas en inglés, Address Resolution Protocol).

Como se puede observar, existen una serie de estados que deben tener una probabilidad asociada de que lleguen a presentarse para entonces, con una cierta certeza, poder inferir que se trata de alguno de los dos tipos de ataques que se tienen identificados en la red bayesiana.

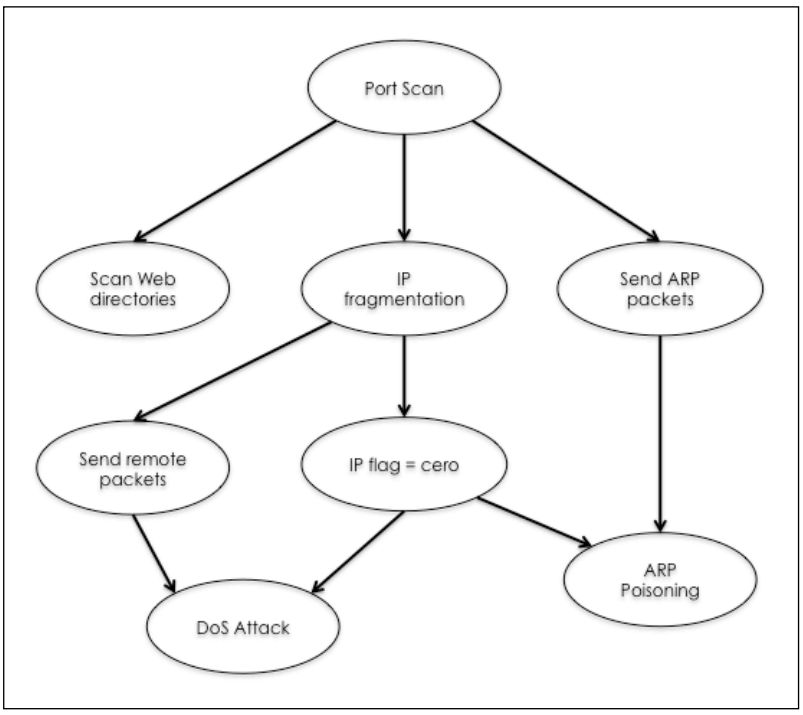

Figura 2.3.5 Ejemplo de una red bayesiana para identificación de ataques informáticos

Podemos observar las suposiciones de independencia condicionada en la red bayesiana (Figura 2.3.5), donde *ARP Poisoning* es condicionalmente independiente de *Port Scan, Scan Web Directories, Send remote packets, IP fragmentation, DoS Attack*. En notación de probabilidad bayesiana quedaría como:

$$P\left(ARPP \middle| PS, SWD, SRP, IPF, DoSA, IPFZ\right) = P\left(ARPP \middle| IPFZ\right) \tag{2.3.8}$$

Donde ARRP es ARP Poisoning, PS es Port Scan, SWD es Scan Web Directories, IPF es IP fragmentation, DoSA es DoS Attack e IPFZ es IP flag = zero.

El aprendizaje de las redes bayesianas está basado en la determinación de dos partes:

Aprendizaje estructural: Alcanzar la identificación de la topología de la red bayesiana.

Aprendizaje paramétrico: Una vez conocida la topología del grafo, se obtienen las probabilidades de cada nodo.

En el aprendizaje paramétrico, se puede tener el caso de que todas la variables sean observables, es decir, el conjunto de datos para el entrenamiento están completos, por lo que regularmente se utiliza el método conocido como *Estimador de máxima verosimilitud*, el cual calcula las probabilidades deseadas partiendo de la frecuencia de los valores provenientes del conjunto de datos del entrenamiento.

Respecto al índice de dispersión (calidad) de las estimaciones, estará en función directa de la cantidad de datos que se tenga en la muestra de entrenamiento. Cuando la muestra no es suficiente, entonces se recurre a la adopción de una distribución de probabilidad para la determinación de las incertidumbres.

La elección del tipo de distribución deberá basarse en pruebas de bondad de ajuste no paramétricas entre dos distribuciones, la de la muestra de datos para entrenamiento y la de las distribuciones de probabilidad más conocidas, para ello se recomienda utilizar los algoritmos de Kolmogorov-Smirnov y Anderson-Darling.

Finalmente el objetivo del aprendizaje paramétrico es encontrar los parámetros asociados a la estructura de la red bayesiana en cuestión. Estos parámetros son las probabilidades a priori de los nodos raíz y las probabilidades condicionales del resto de las variables considerando como antecedentes a los nodos padres.

Respecto al aprendizaje estructural, su objetivo es la identificación de la estructura de la red, tarea para la que el grado de dificultad está en función del número de nodos, y restricciones entre ellos, de tal suerte, que el aprendizaje estructural de una red con gran cantidad de nodos y sin restricciones puede parecer poco menos que imposible puesto que número de grafos que pueden derivarse es ilimitado.

Aun así existen algoritmos que atienden el aprendizaje de redes sin restricciones, los cuales aplican métodos para explorar las relaciones de dependencia entre conjuntos de nodos comenzando en pares hasta formar estructuras de exploración parecidas a clúster, que podrían dar luz a la forma en la que se puede determinar su conexión entre sí.

Un algoritmo utilizado para este tipo de casos es el llamado TAN (Tree Augmented Naive-Bayes), que es un clasificador bayesiano simple aumentado con un

árbol. Este algoritmo parte del supuesto de que se tienen restricciones en la estructura de la red bayesiana. La restricción base de este modelo es que los nodos atributos formen solamente un árbol.

Cuando no se imponen restricciones sobre la estructura de las dependencias entre los atributos de la red bayesiana, se utiliza el clasificador llamado BAN (Bayesian Network Augmented Naive-Bayes), que permite construir cualquier grafo para la red bayesiana a diferencia del TAN que se restringe a solamente construir árboles. (Friedman, Geiger, & Goldszmidt, 1997; Pourret, Naïm, & Marcot, 2010)

Por supuesto que realizar el aprendizaje estructural a través del clasificador bayesianas BAN, es tan complejo como construir la red bayesiana sin considerar ningún nodo clase.

El clasificador bayesiano TAN es mucho más utilizado para conjuntos pequeños de nodos de una red bayesiana amplia, por su naturaleza de aprendizaje restringido a árboles.

Cuando se ha determinado implantar el aprendizaje estructural sin restricciones, se pueden utilizar algoritmos de búsqueda heurística, pero cuando la cantidad de nodos es amplia, resulta profundamente lenta e ineficiente, por lo que se tiene que combinar con otras técnicas que permita reducir el espacio de posibilidades de creación de grafos, como pueden ser algoritmos genéticos, búsquedas bidireccionales, entre varias otras.

También es posible utilizar algoritmos globales de ajuste, entre los más conocidos están las medidas bayesianas y el de mínima longitud de descripción.

Las medidas bayesianas tienen como propósito maximizar la probabilidad de la estructura de la red basado en los datos de entrenamiento que se estén considerando, el principio es comparar el valor obtenido para distintas estructuras.

El algoritmo de mínima longitud de descripción MDL (por sus siglas en inglés, Minimum Description Length) está en la familia de los métodos basados en calificaciones, Score-based Methods (Pourret, Naïm, & Marcot, 2010). El proceso asigna una calificación (score) a cada red bayesiana que se considera candidato, típicamente una

de las medidas que se toman en cuenta es que tan correctamente la red bayesiana es descrita por el conjunto de datos *D*. Considerando una estructura *G*, su calificación puede expresarse como sigue:

$$Score(G,D) = \mathrm{P}(G|D) \tag{2.3.9}$$

Esto indica que la probabilidad posterior de *G* dado el conjunto de datos. Este es un algoritmo basado en calificaciones que intenta maximizar la calificación expresado en términos de la ley de Bayes:

$$Score(G,D) = P(G|D) = \frac{P(D|G)P(G)}{P(D)} \tag{2.3.10}$$

Para maximizar la ecuación 2.3.10, tendríamos que maximizar únicamente el numerador, puesto que el denominador no está en función de *G*. Para calcular P(D | G) podemos utilizar la propuesta de Heckermand (1995), Rebane & Pearl (1987) y Friedman, Gieger & Goldszmidt (1998) en las que realizan una aproximación estandarizada sobre todos los posibles parámetros, ponderándolos cada uno a través de la determinación de su probabilidad posterior:

$$P(D|G) = \int P(D|G,p)P(p|G)dp \tag{2.3.11}$$

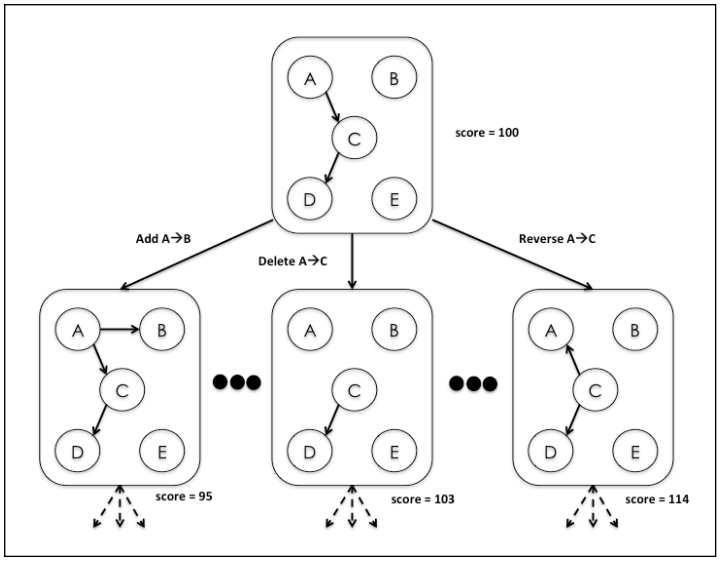

Figura 2.3.6 Ejemplo de la estructura de una red bayesiana por el procedimiento de búsqueda "Hill-climbing" (Margaritis, 2003)

3. DISEÑO METODOLÓGICO DE LA INVESTIGACIÓN

Introducción

En este capítulo se ven los elementos metodológicos que se aplicaron para el diseño de la investigación desde la óptica de la metodología, la presentación de las hipótesis y variables de la investigación, la conformación de la investigación exploratoria y cuantitativa, así como sus alcances y limitaciones, con el fin de determinar de forma clara y sencilla el ámbito de acción del trabajo, que dan sustento a las hipótesis de la investigación.

3.1. Hipótesis

Las hipótesis de investigación que se abordan en este trabajo son las siguientes:

- Es posible identificar un ataque por DDoS a una red de datos a través de un modelo basado en una red bayesiana que permita el diagnóstico de anomalías en el tráfico.

- Es posible mejorar la calidad de la información recolectada para análisis forense a través de la implantación de un modelo de red bayesiana.

3.2. Diseño de la investigación

Este trabajo comenzó con una investigación exploratoria sobre trabajos realizados en los últimos 5 años sobre el tema del análisis forense y la identificación de incidencias y ataques sobre redes de datos.

Esto permitió la identificación del marco teórico sustento de la presente investigación, particularmente sobre la teoría de las redes bayesianas y sus aplicaciones así como las implantaciones de análisis forense e identificación de incidencias y ataques sobre redes de datos.

Posteriormente se identificaron las principales características del ataque conocido como DoS para poder realizar el diseño de la red bayesiana que permita realizar inferencias basados en los parámetros del diseño.

Respecto a la elección del tipo de método para la identificación de evidencia sobre los incidentes o anomalías, existen varios métodos de aprendizaje autónomo, para atender este tipo de requerimientos, como son las redes neuronales y las redes bayesianas, ambos métodos ofrecen una grupo de características que son ampliamente explotadas dependiendo de la aplicación y el entorno en que se implanten.

Esto nos permite ampliar los campos de aplicación de las redes bayesianas y las neuronales en la modelación de problemas y entornos muy diversos.

En la ciencia forense, las redes bayesianas han sido un método recurrente de utilización, por las bondades que ofrece respecto a la representación e interpretación del entorno, haciéndolo sencillo y amigable para comprensión y explicación.

Para modelado del tiempo de respuesta de sistemas orientados a servicio las redes neuronales, ofrecen rapidez en la evaluación y generación de predicciones, por utilizar pequeños conjuntos de datos y rutinas que no requieren una gran cantidad de parametrizaciones. (Zhang & Bivens, 2007; Correa, Bielsa, Teixeira & Alique, 2006)

Las redes neuronales son mucho más eficiente cuando se cuenta con una gran cantidad de datos y alta incertidumbre en la forma en que se han producido estos, es decir, un poco conocimiento de la estructura y los parámetros que modelan el problema.

Las redes bayesianas son menos sensibles al tamaño del conjunto de datos por lo que son más propicias para entornos que tienden a cambios constantes y de forma rápida, con moderada o poca cantidad de datos, lo que lleva a la necesidad de reconstrucción del modelo frecuentemente.

Considerando que el análisis forense se realiza *post mortem*, el tiempo de aproximación inferencial, no resulta un factor relevante para , que las redes bayesianas no cumplan los requerimientos necesarios para poder ofrecer una aproximación a la aportación de argumentos sobre las evidencias que se evalúan.

Las redes bayesianas han demostrado un desempeño eficiente cuando atienden problemas relacionados con un gran número de pequeñas hipótesis que deben irse evaluando no de forma individual, sino en conjunto o en relación con la certeza o no de

las otras hipótesis, que puedan ir ofreciéndose como evidencias, es decir, realizar análisis en conjunto de forma interdependiente, donde el teorema de Bayes emerge como una opción robusta para atender este tipo de modelos.

Se determinó la utilización de redes bayesianas, porque existen investigaciones recientes donde se demuestra que éstos métodos de aprendizaje automático pueden ser útiles para la toma de decisiones autónomas, así como para la comprensión del comportamiento del sistema que se está modelando. (Zhang & Bivens, 2007; Taroni, Aitken, Gabrolino & Biedermann, 2010; Rekhis, Krichene, Boudriga, 2009; Taroni, Bozza, & Biedermann, 2005; Pourret, Naïm, & Marcot, 2010)

La utilización de redes bayesianas para la identificación de incidencias, anomalías o ataques en una red de datos, tiene como principio fundamental el ofrecerle a los investigadores forenses herramientas que les ayuden a captar de forma selectiva e inteligente información para su análisis y esto permita, reducir el time-gap entre el ataque y el análisis forense, con la finalidad de conocer mejor al atacante y sus técnicas.

Una vez realizado el diseño, se continuó con la prueba de la red bayesiana a través de la recolección de tráfico de una red, su filtrado para evitar ataques obvios y entonces realizar la inferencia de la red bayesiana, la cual permitió realizar ajustes a los parámetros definidos en el diseño y entonces obtener las métricas definitivas.

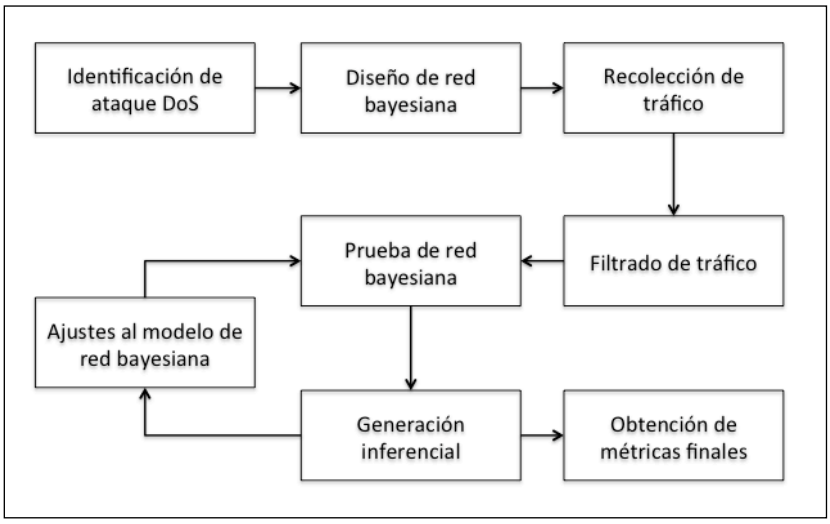

Figura 3.2.1 Diagrama del diseño cuasi-experimental de la investigación con retroalimentación para ajuste

En la Figura 3.2.1 se puede observar el diagrama del diseño cuasi-experimental que se utilizó para la presente investigación, el cual incluye, una retroalimentación después de la generación inferencial de la red bayesiana para efecto de ajustar sus parámetros que se definieron en el diseño de la misma.

3.3. Alcances y limitaciones

Este trabajo de investigación tuvo como alcance el diseño de una red bayesiana que permita identificar incidentes en una red de datos de un ataque en específico.

Así mismo fue probada su efectividad a través de la identificación de un ataque conocido como DoS (por sus siglas en inglés, Denial of Services).

Este trabajo se limitó a la identificación de un ataque conocido modelado a través de las técnicas de redes bayesianas, las cuales permiten ofrecer modelos viables para el análisis forense de cualquier disciplina.

3.4. Viabilidad

Se consideró que la viabilidad de esta investigación fue alta, toda vez que se basó en modelos estadísticos probados en el área de conocimiento del análisis forense así como la utilización de herramientas y modelos utilizados en la industria de la seguridad de información que pueden estar al alcance de cualquier empresa o institución educativa que desee replicar los modelos propuestos.

3.5. Aportaciones

Se considera que el presente trabajo tiene las siguientes aportaciones:

Revisar y elaborar un compendio sobre el marco teórico y estado del arte del análisis forense y las redes bayesianas.

La posibilidad de modelar cualquier incidente que puede presentarse en una red de datos, que permita dar inicio a la recolección de información completa e integral en tiempo real.

Proponer a las redes bayesianas como un modelo viable para la identificación de posibles ataques en red que permita optimizar los dispositivos de seguridad y recolección de información.

4. MODELO DE UNA RED BAYESIANA PARA IDENTIFICACIÓN DE ATAQUES POR DDOS

Introducción

En este capítulo se mostrará el modelo desarrollado para la identificación de incidencias y ataques a una red de datos a través de una red bayesiana.

Así mismo se presentarán los resultados obtenidos en la prueba realizada en un prototipo de una red bayesiana para la detección de un ataque de DDoS.

4.1. Diseño

El diseño de la red bayesiana se basó inicialmente en la obtención del tráfico de datos en una red, filtrado a través de un IDS Snort, permitiendo eliminar aquellos ataques que ya son identificados por estas herramientas, para ser inyectado éste tráfico en tiempo real a una red bayesiana, que fue diseñada para la identificación específica de ataques DDoS.

Una vez que la red bayesiana recibe el tráfico filtrado, comienza a realizar sus inferencias para determinar si encuentra evidencia suficiente para pronosticar un ataque por DDoS. En caso de encontrar tal evidencia, entonces lanza una alerta para indicar la incidencia identificada y su nivel de probabilidad de ocurrencia con un intervalo de confianza del 95%.

En la Figura 4.1.1 se puede observar el diseño que se propuso en este trabajo de investigación para la identificación de incidencias a través de una red bayesiana. Como parte del alcance de este trabajo, se determinó el modelo de la red bayesiana para la identificación de un ataque de DDoS, sin embargo, se puede modelar cualquier tipo de ataque conocido o los elementos mínimos de evidencia, que permitan a la red bayesiana suponer algún ataque no conocido.

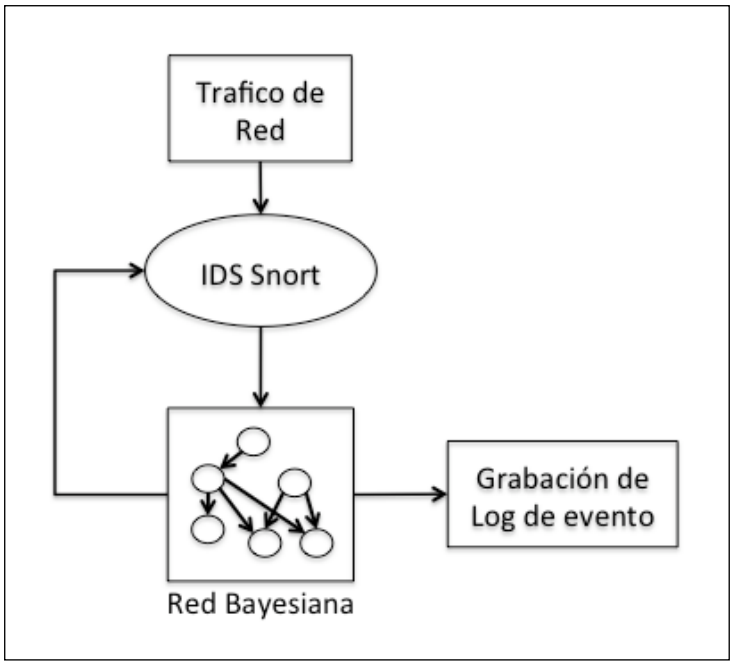

Figura 4.1.1 Diagrama del diseño de la identificación de un ataque a través de una red bayesiana

En la Figura 4.1.2 se puede observa el diseño extendido del trabajo de identificación de incidencias en una red de datos a través de una red bayesiana, que arranque en tiempo real, la recolección de evidencia a través de una HoneyNet.

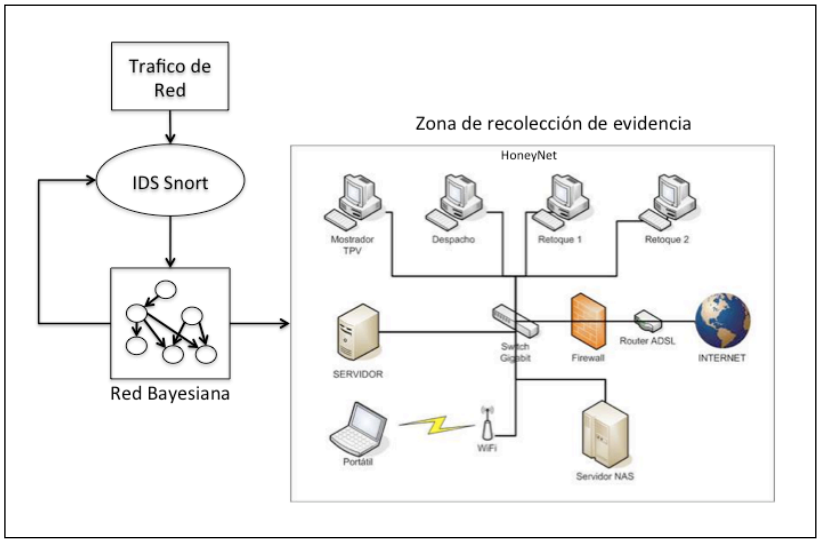

Figura 4.1.2 Diagrama del diseño extendido de la identificación de un ataque a través de una red bayesiana para inciar la recolección de evidencia a través de una HoneyNet.

La incidencia específica con la que se diseñó la red bayesiana fue el ataque conocido como DDoS, que desde el año 2000 el US-CERT lo ha identificado como el ataque con mayor incidencia en Internet con objetivos ubicados en los Estados Unidos.

(Bursh & Bauer, 2000) presentan la inferencia de una ruta de ataque por inundación de todas las ligas con una gran explosión de tráfico en la red, la cual se encargó de medir e identificar.

(Stone, 2000) en su investigación propone la creación de un centro de rastreo que automatice el mecanismo de búsqueda infiriendo la ruta del ataque a través del ruteo que haya seguido el tráfico de la red.

(Belloving, 2000) propone la auditoria de paquetes en la víctima del ataque, en su canal de mensajes de ICMP.

La utilización explícita del soporte a redes de datos fue propuesto por Snoren, 2002, donde la bitácora de auditoria de los paquetes, se almacenara de forma distributiva en la misma red de datos.

En el trabajo de Belenky & Ansari, 2003 se propone el uso de una técnica de señalización determinista de paquetes que viajan en la red y una versión de este trabajo optimizando la ubicación local de cada una de las señalizaciones fue propuesta por Al-Duwairi & Daniels, 2004 y por Muthuprasanna & Manimaran, 2005.

Savage, Wetherall, Karlin & Anderson, 2000 propusieron una señalización probabilística de los paquetes que viajan en la red, empleando una representación de nodos y conexiones reduciendo los tamaños de los bits por paquete mientras se expanden los fragmentos probabilísticamente a través de múltiples paquetes.

El trabajo de Song & Perrig, 2001 retoman el modelo probabilístico añadiéndole un esquema de seguridad utilizando una técnica de autenticación para la integridad de los datos que viajan en la red.

(Dean et. al., 2000) propusieron un esquema nuevo de codificación utilizando una enfoque algebraico para la integración de la información de la ruta de ataque.

De los múltiples mecanismos que se han desarrollado para la detección de anomalías originadas por los ataques de DDoS, se determinó utilizar dos modelos que presentaron una tasa alta de efectividad y una significativa tasa baja de falsos positivos que en conjunto y modeladas por una red bayesiana, puede ofrecer un modelo creativo y eficiente.

Para la construcción de la red bayesiana se utilizó el modelo llamado "Detección de DDoS basado en la densidad de conexión de una vía" de Xu, He & Lou, 2007 y el modelo conocido como "Detección de DDoS utilizando las características de interacción de las direcciones IP's" por Cheng, Yin, Cai & Wu, 2009; identificando las variables aleatorias de cada modelo y sus dependencias condicionales para la construcción de los parámetros probabilísticos de la red bayesiana.

4.2. Evaluación del modelo

Introducción

En esta parte se mostrarán las condiciones bajo las cuales se realizó la evaluación del modelo de la red bayesiana para la identificación de anomalías en una red de datos, particularmente un ataque por DDoS y verificar los valores de las variables aleatorias del modelo para su afinación y parametrización.

4.2.1. Definición de la muestra

Basado en los trabajos de Xu, He & Lou, 2007 y de Cheng, Yin, Cai & Wu, 2009, se pudieron identificar los vectores de las variables aleatorias necesarias para la estimación de las distribuciones de probabilidad de cada una de ellas en las franjas de tiempo establecidas.

Para posteriormente, comenzar el análisis estadístico que nos permitió identificar las probabilidades de dependencia e independencia entre las variables aleatorias que fueron suministradas a la red bayesiana propuesta.

Los vectores de variables almacenadas para su análisis estadístico en fracciones de 5 segundos, identificados en T_i, donde *i:1,2,3…,n*; fueron:

- *Interaction Flow* (**IF**): Interacción del Flow de paquetes.
- *Timestamp*: Identificador del tiempo de la muestra.
- *IP Address Source* (**IPS**): Dirección IP del nodo origen.
- *IP Address Destination* (**IPD**): Dirección IP del nodo destino
- *Classific of Packets* (**C_Pckts**)*:* Clasificación de paquetes en IPS o IPD
- *Number of Destination Port* (**NDP**): Número de Puerto del nodo destino.
- *Source Half Interaction Flow* (**SHIF**): Flujo con solo IPS sin IPD.
- *Distribution of Source IP Address* (**DSIPa**): Distribución de paquetes con la misma dirección IP Origen.
- *Concentration of Destination IP Address* (**CDIPa**): Concentración de paquetes con la misma IP Destino.
- *Outburst in Network Traffic* (**ONT**): Explosión de tráfico en la red.
- *Protocol Flags* (**PF**): Banderas de cada paquete de cada protocolo.
- *One Packets Connection* (**OPC**): Paquetes de datos enviados y/o recibidos sin un par confirmado.
- *Two Packets Connection* (**TPC**): Paquetes de datos enviados a un destino y con respuesta del origen/destino confirmado.
- *Density One Packets Connection* (**DOPC**): Densidad de paquetes con una sola conexión confirmada.
- *Length Average IP Flow* (**LAIPF**): Tamaño promedio del flujo de paquetes IP's.

- *Ratio of Incoming and Outgoing Packets* (**RIOP**): Razón de paquetes IP's de entrada y salida.
- *Entropy of Length Average IP Flow* (**E_LAIPF**): Aleatoriedad de los tamaños promedio de los flujos de paquetes IP's.
- *Entropy of Protocols* (**EP**): Aleatoriedad de los protocolos de los paquetes recibidos (TCP, UDP, ICMP).
- *IP Address same Source* (**IPASS**): Paquetes con la misma dirección IP origen.
- *IP Address same Destination* (**IPASD**): Paquetes con la misma dirección IP destino.
- *Source Half Interaction Flow* (**SHIF**): Paquetes con interacción entre nodos intermedios.
- *Outburst in Network Traffic* (**ONT**): Nivel de explosión del tráfico de la red.

De cada uno de los vectores de información recopilados a través del software de análisis de protocolos de red Wireshark y almacenados en una base de datos SQL Server, se realizaron los análisis estadísticos para efecto de generar las probabilidades *a priori*, que serían alimentadas al modelo de la red bayesiana propuesta.

4.2.2. Recopilación de información

Se construyó una base de datos almacenando las principales variables del tráfico de una red experimental formada por un servidor DELL Optiplex GX520, Pentium 4HT, 2.00 GHz, 2GB Memoria RAM con Windows Server 2003 SP3, 3 Desktops de usuario DELL Vostro 230 Slim Tower, Intel Celeron 450, 2.20GHz, 2GB RAM, Windows Vista

SP3 y 2 Desktop atacates DELL Vostro 230 Slim Tower, Interl Core2Duo E7500, 2.93GHz, 4GB RAM, Windows 7 Professional 64bits.

En la Figura 4.1.3 se muestra un esquema gráfico de la topología de red que fue utilizada para la elaboración de la valoración del modelo de la red bayesiana.

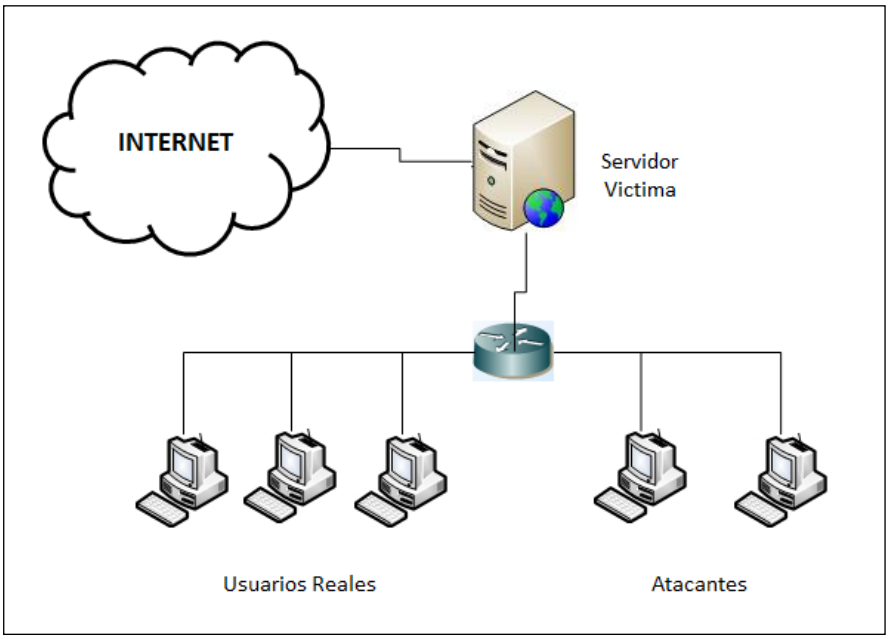

Figura 4.1.3 Topología de la red experimental.

Los ataques por DDoS se realizaron a través de las herramientas jolt2.exe fragmentador de paquetes bajo UDP y ICMP, fragroute.exe que es un fragmentador de paquetes bajo TCP, Juno-Z.101f.C es un amplificador de paquetes TCP SYN y HGod.exe que es un generador de inundaciones de paquetes para ambientes Windows.

En las Figuras 4.1.4 a la 4.1.7 se presentan las distribuciones de probabilidad de frecuencia *pdf* (por sus siglas en inglés, probability distribution frecuency) de algunas variables aleatorias considerando un flujo normal de paquetes y flujo bajo ataque.

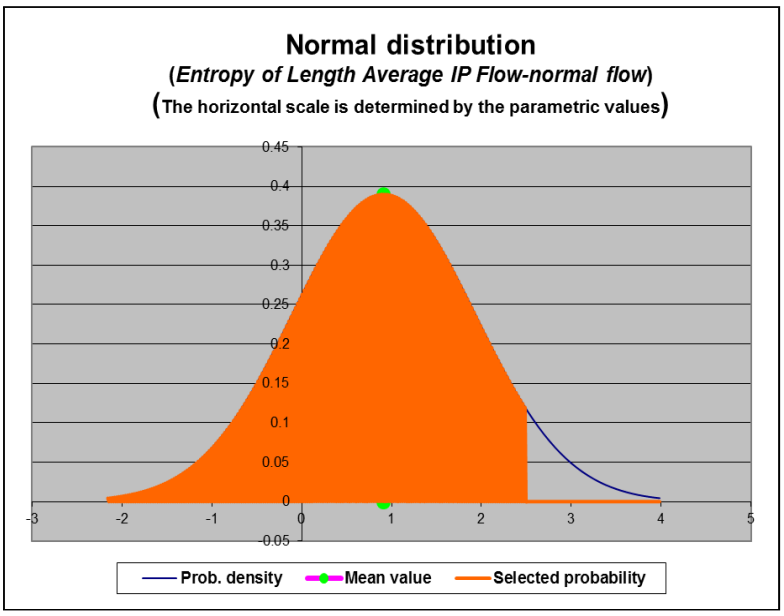

Figura 4.1.4 Gráfico de la distribución de probabilidad de la variable "Entropy of Length Average IP Flow" bajo un flujo normal.

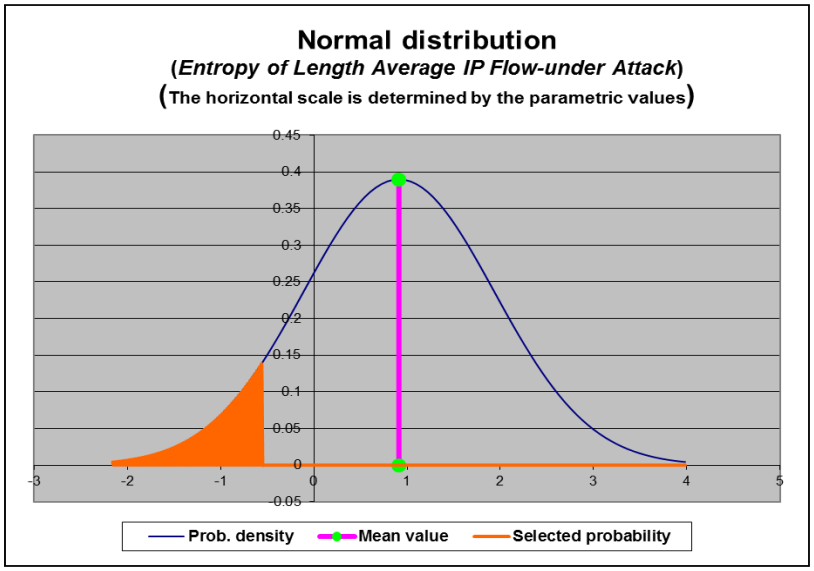

Figura 4.1.5 Gráfico de la distribución de probabilidad de la variable "Entropy of Length Average IP Flow" bajo un flujo de ataque de DDoS.

Como se puede observar, la variable aleatoria tiene una distribución de probabilidad normal y ofrece basado en el análisis estadístico a la base de datos experimental, que la probabilidad de tener niveles de entropía bajos con flujos de paquetes caracterizados como tráfico regular, es de un 83.192% y alcanzar niveles bajos de entropía bajo flujos considerados como ataques es de 16.808%.

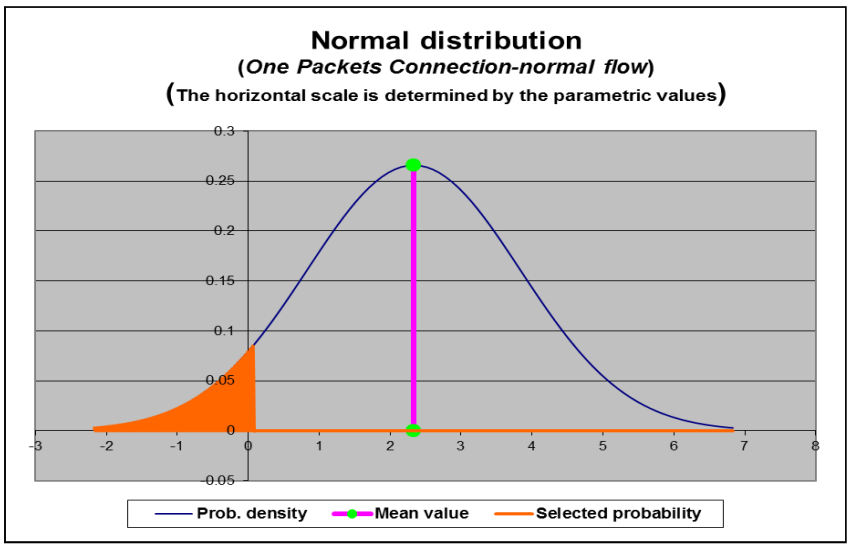

Figura 4.1.6 Gráfico de la distribución de probabilidad de la variable "One Packet Connection" bajo un flujo normal.

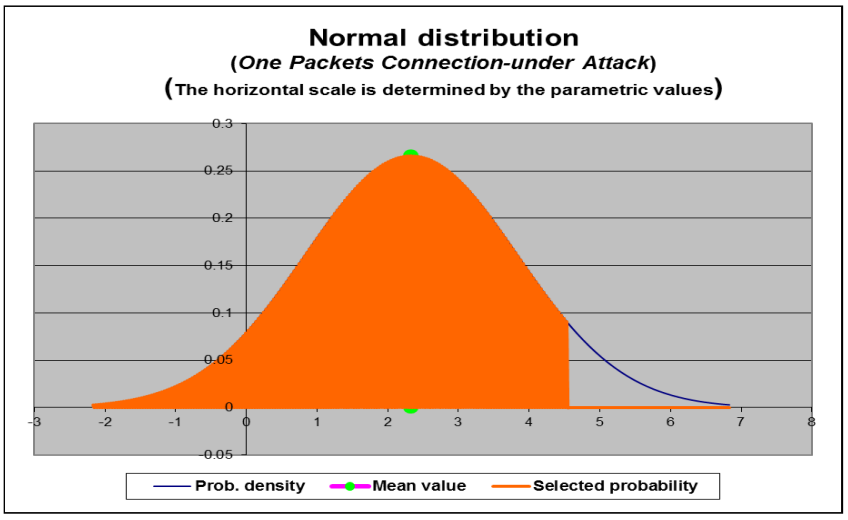

Figura 4.1.7 Gráfico de la distribución de probabilidad de la variable "One Packet Connection" bajo un flujo de ataque de DDoS.

Como se puede observar, la variable aleatoria tiene una distribución de probabilidad normal y ofrece basado en el análisis estadístico a la base de datos experimental, que la probabilidad de tener paquetes con una sola conexión bajo flujos caracterizados como tráfico regular, es de un 6.837% y la probabilidad de tener paquetes con una sola conexión bajo flujos considerados como ataques es de 93.163%.

De igual forma, se realizaron el análisis de las variables restantes así como de algunas adicionales que fueron necesarias incluir en la red bayesiana una vez que se revisó su comportamiento inicial.

4.3. Resultados del modelo

Introducción

En esta sección se presentan el modelo de la red bayesiana con las probabilidades a priori estimadas a través del análisis estadístico de la base de datos obtenida durante el modelo experimental de 11 horas diarias de tráfico de información durante 7 días con franjas de recolección de los vectores de variables aleatorias de cada 5 segundos.

4.3.1. Descripción de la red bayesiana

La construcción de la red bayesiana, se determinó bajo la identificación inicial de las variables aleatorias que determinan el modelo propuesto de detección de anomalías en el tráfico de una red de datos.

Una vez que se logró decidir las variables aleatorias del modelo, se realizó la clasificación de las mismas sobre la estructura típica causal de una red bayesiana, el cual se muestra en la Figura 4.1.8.

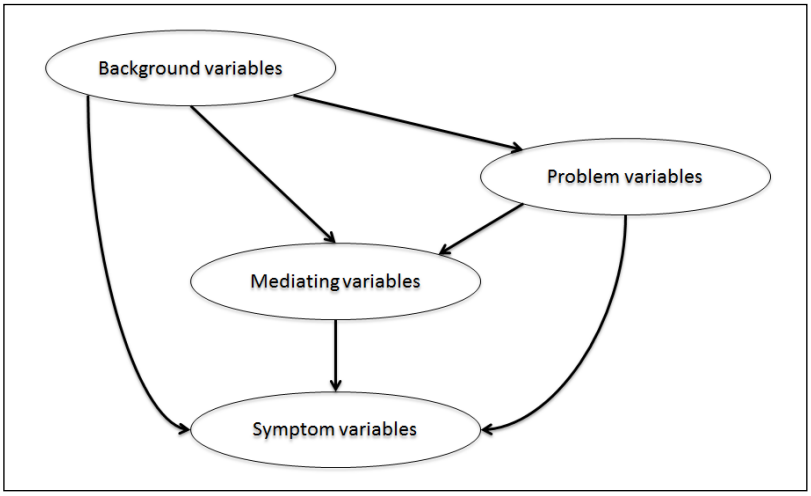

Figura 4.1.8 Típica estructura causal de una red bayesiana (Kjaerulff & Madsen, 2008).

Después de la clasificación de las variables aleatorias, se verificó la estructura resultante contra los perfiles típicos de las redes bayesiana, para corroborar su correcta determinación y establecimiento de las relaciones de dependencia. Véase la Figura 4.1.9.

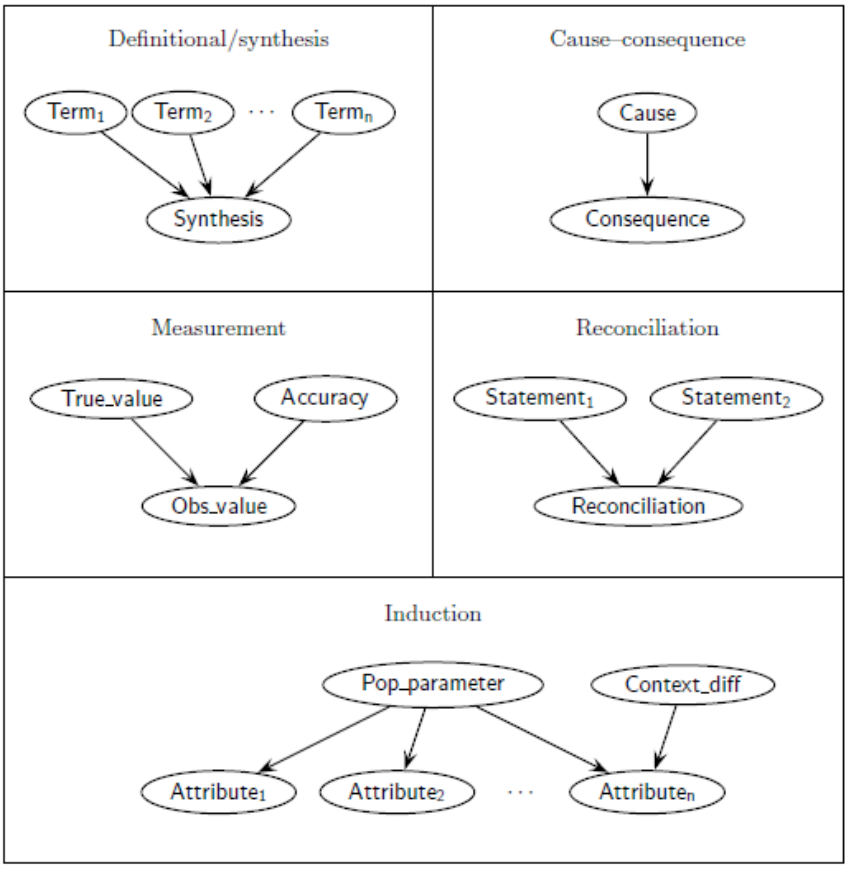

Figura 4.1.9 Los cinco tipos básicos de estructuras causales para redes bayesianas (Kjaerulff & Madsen, 2008).

La estructura resultante del modelo de la red bayesiana para la identificación de anomalías en una red datos originados por un ataque de DDoS se muestra en la Figura 4.1.10.

Como se puede observar, esta red bayesiana tiene una combinación de dos estructuras básicas conocidas como *cause-consequense* y *definitional/synthesis*, para determinar la identificación de un ataque de DDoS.

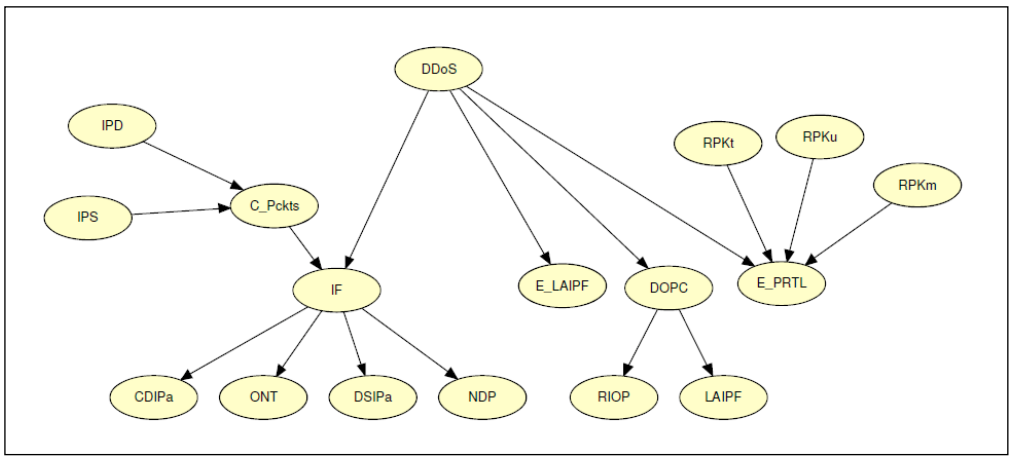

Figura 4.1.10 Modelo de la Red Bayesiana propuesta para la identificación de un ataque de DDoS.

La variable aleatoria DDoS se identificó como una variable tipo *consecuencia* que deberá ser determinada a través de la resolución de las dependencias condicionales de toda la red bayesiana propuesta.

El mini-árbol formado a partir del nodo IF (Interaction Flow), es la red construida basada en el modelo de detección utilizando las características de interactividad de las direcciones IP's del tráfico analizado.

El mini-árbol formado de por los nodos E_LAIPF, DOPC y E_PRTL son basados en el modelo de detección considerando las características de los paquetes del tráfico analizado respecto a las conexiones incompletas entre origen y destino.

Todas las variables aleatorias están en función del tiempo *t* que es registrado en franjas de tiempo *i* que va desde *i=1,2,3…,n* y es analizado por la cantidad de paquetes *j* que van desde *j=1,2,3…,n* de los vectores *m* que van desde *m=1,2,3,…,n*.

La determinación de la variable DOPC (*Density One Packets Connection)* que está en función del tiempo *t*, está dada por:

$$DOPC(t) = \frac{\sum OPC(t)}{\sum Pk(t)} \tag{4.3.1}$$

La variable LAIPF (Length of IP Flow) está definida por:

$$LAIPF = \frac{\sum IPpkts}{\sum IPflows} \tag{4.3.2}$$

Considerando que los *IPflows* se definen como un conjunto de paquetes que tienen los mismo cinco elementos de un paquete coincidentes (*IP Source, Port Source, IP Destination, Port Destination y Protocol*). Los *IPpkts* son el conjunto de paquetes recibidos y almacenados para su análisis.

La variable RIOP (*Ratio of Incoming and Outgoing Packets*) está definida por:

$$RIOP = \frac{\sum Incoming\ IPpkts}{\sum Outgoing\ IPpkts} \tag{4.3.3}$$

La variable E_LAIPF (*Entropy of Length Average IP Flow)* está definida por:

$$E_{LAIPF} = -\omega_j \sum_{j=1}^{n} log_2\ \omega_j \tag{4.3.4}$$

Dónde:

$$\omega_j = \frac{LAIPF_j}{\sum IPpkts}$$

La variable E_PRTL (*Entropy of Protocols)* está definida por:

$$E_{PRTL} = -Rpkt_t log_2 Rpkt_t - Rpkt_u log_2 Rpkt_u - Rpkt_m log_2 Rpkt_m \tag{4.3.5}$$

Dónde:

$Rpkt_t$ es la tasa de paquetes TCP.
$Rpkt_u$ es la tasa de paquetes UDP y
$Rpkt_m$ es la tasa de paquetes ICMP

La variable IF (*Interaction Flow*) está definida como:

$$IF = \{IPS | IPS \cup IPD \neq IPS; \forall\ IPS, IPD \in IPpkts\} \tag{4.3.6}$$

La variable SHIF (*Source Half Interaction Flow*) está definida como:

$$SHIF = \{IPS | IPS \cup IPD = IPS; \forall\ IPS, IPD \in IPpkts\} \tag{4.3.7}$$

Donde Ippkts son el conjunto completo de paquetes, IPS son el conjunto de paquetes con la misma dirección IP origen e IPD son el conjunto de paquetes con la misma dirección IP destino.

La variable NDP (*Number of Destinarion Port*) está definida como:

$$NDP = \sum_{j=1}^{n} SHIF_j > \theta \tag{4.3.8}$$

Donde θ es el umbral de un rango de número de NDP's por milisegundo.

La variable DSIPa (*Density Same IP Address*) está definida como:

$$DSIPa = \frac{\sum SHIF_s - \sum SHIF_d}{\sum IF} > 1 \tag{4.3.9}$$

Donde $SHIF_s$ son el conjunto de paquetes SHIF que tiene la misma dirección IP origen y $SHIF_d$ son el conjunto de paquetes SHIF que tienen la misma dirección IP destino.

La variable CDIPa (Count Density IP Address) está definida como:

$$CDIPa = \sum SHIF_d \rightarrow 0 \tag{4.3.10}$$

La variable ONT (*Outburst in Network Traffic*) está definida como:

$$ONT = \sum SHIF_s - \sum SHIF_d > \beta \tag{4.3.11}$$

Donde β es el umbral de un rango de número de paquetes por milisegundo del tipo SHIF el cual tenderá a ser muy grande en el tiempo *t* en condiciones de un ataque.

4.3.2. Tratamiento a la red bayesiana

La red bayesiana una vez que fue definida bajo la topología demostrada en la Figura 4.1.10, se le realizó la primera carga de parametrización de las variables aleatorias basadas en el análisis estadístico realizado a la base de datos que contiene el tráfico de una red experimental para obtener los valores *a priori*.

La parametrización a la red bayesiana se realizó para todos y cada uno de los nodos definidos en la primera topología y se utilizó el modelo Naive Bayes con

optimización triangular para la generación de la primera inferencia paramétrica de la red bayesiana, basado en el siguiente modelo:

Optimización Triangular para Naive Bayes:

$$OpTrgNB = \frac{\left[a - ln(w(s)) + b\sqrt{\frac{n}{n-1}f(s)}\right]^c}{\Delta} \quad (4.3.12)$$

Se consideraron los valores de *a=1.0, b=0.1, c=3.0 y separadores máximos= 100,000*, utilizando la *propagación automática*, bajo la aproximación inferencial inicial que se propone para el primer modelo de una red bayesiana por Darwiche, 2009; con no más de 18 nodos y 30 relaciones.

Los valores de aproximación inicial de cada nodo de la red bayesiana se pueden observar en las tablas siguientes:

IF

DDoS	Attack		Normal	
C_Pckts	IF	SHIF	IF	SHIF
Small	0.8	0.2	0.2	0.8
Large	0.2	0.8	0.8	0.2

C_Pckts

IPD	Empty		No Empty	
IPS	Empty	No Empty	Empty	No Empty
IF	1	0	1	1
SHIF	0	1	0	0

ONT

IF	Small	Large
Small	0.2	0.8
Large	0.8	0.2

CDIPa

IF	Small	Large
Small	0.8	0.2
Large	0.2	0.8

LAIPF

DOPC	<=0.30	>0.30
<= 1	0.44	0.56
> 1	0.58	0.42

E_LAIPF

DDoS	Attack	Normal
<=7	0.09	0.9
>7	0.91	0.1

NDP

IF	Small	Large
>= threshold	0.8	0.2
< threshold	0.2	0.8

DSIPa

IF	Small	Large
Small	0.2	0.8
Large	0.8	0.2

IPS

Empty	0.5
No Empty	0.5

IPD

Empty	0.5
No Empty	0.5

DOPC

DDoS	Attack	Normal
<=0.30	0.14	0.81
>0.30	0.86	0.19

RIOP

DOPC	<=0.30	>0.30
<=500	0.46	0.54
>500	0.44	0.56

DDoS

Attack	0.5
Normal	0.5

E_PRTL

DDoS	Attack									
RPKm	High									Medium
RPKu	High			Medium			Low			High
RPKt	High	Medium	Low	High	Medium	Low	High	Medium	Low	High
<=0.21	0.673469	0.795918	0.714286	0.591837	0.632653	0.77551	0.734694	0.55102	0.734694	0.734694
>0.21	0.326531	0.204082	0.285714	0.408163	0.367347	0.22449	0.265306	0.44898	0.265306	0.265306

DDoS	Attack									
RPKm	Medium								Low	
RPKu	High			Medium			Low		High	
RPKt	Medium	Low	High	Medium	Low	High	Medium	Low	High	Medium
<=0.21	0.571429	0.673469	0.55102	0.55102	0.693878	0.673469	0.571429	0.714286	0.795918	0.755102
>0.21	0.428571	0.326531	0.44898	0.44898	0.306122	0.326531	0.428571	0.285714	0.204082	0.244898

DDoS	Attack							Normal		
RPKm	Low							High		
RPKu	High	Medium			Low			High		
RPKt	Low	High	Medium	Low	High	Medium	Low	High	Medium	Low
<=0.21	0.857143	0.755102	0.612245	0.918367	0.836735	0.571429	0.510204	0.673469	0.795918	0.714286
>0.21	0.142857	0.244898	0.387755	0.081633	0.163265	0.428571	0.489796	0.326531	0.204082	0.285714

DDoS	Normal									
RPKm	High						Medium			
RPKu	Medium			Low			High			Medium
RPKt	High	Medium	Low	High	Medium	Low	High	Medium	Low	High
<=0.21	0.591837	0.632653	0.77551	0.734694	0.55102	0.734694	0.734694	0.571429	0.673469	0.55102
>0.21	0.408163	0.367347	0.22449	0.265306	0.44898	0.265306	0.265306	0.428571	0.326531	0.44898

DDoS	Normal									
RPKm	Medium						Low			
RPKu	Medium			Low			High		Medium	
RPKt	Medium	Low	High	Medium	Low	High	Medium	Low	High	Medium
<=0.21	0.55102	0.693878	0.673469	0.571429	0.714286	0.795918	0.755102	0.857143	0.755102	0.612245
>0.21	0.44898	0.306122	0.326531	0.428571	0.285714	0.204082	0.244898	0.142857	0.244898	0.387755

E_PRTL

DDoS	Normal			
RPKm	Low			
RPKu	Medium	Low		
RPKt	Low	High	Medium	Low
<=0.21	0.918367	0.836735	0.571429	0.510204
>0.21	0.081633	0.163265	0.428571	0.489796

RPKu

High	0.05
Medium	0.4
Low	0.55

RPKt

High	0.06
Medium	0.4
Low	0.54

RPKm

High	0.07
Medium	0.4
Low	0.53

Como se puede observar, en el nodo que se considera como la variable aleatoria *consecuencia* que es DDoS, se consideró como parámetro a priori el valor del 0.5 de probabilidad para ambos casos, concediendo el supuesto de una igualdad en la posibilidad de que se genere un ataque.

En la Figura 4.1.11 se puede observar la red bayesiana resultante una vez que se realizó el cálculo de la primera inferencia con los parámetros iniciales definidos.

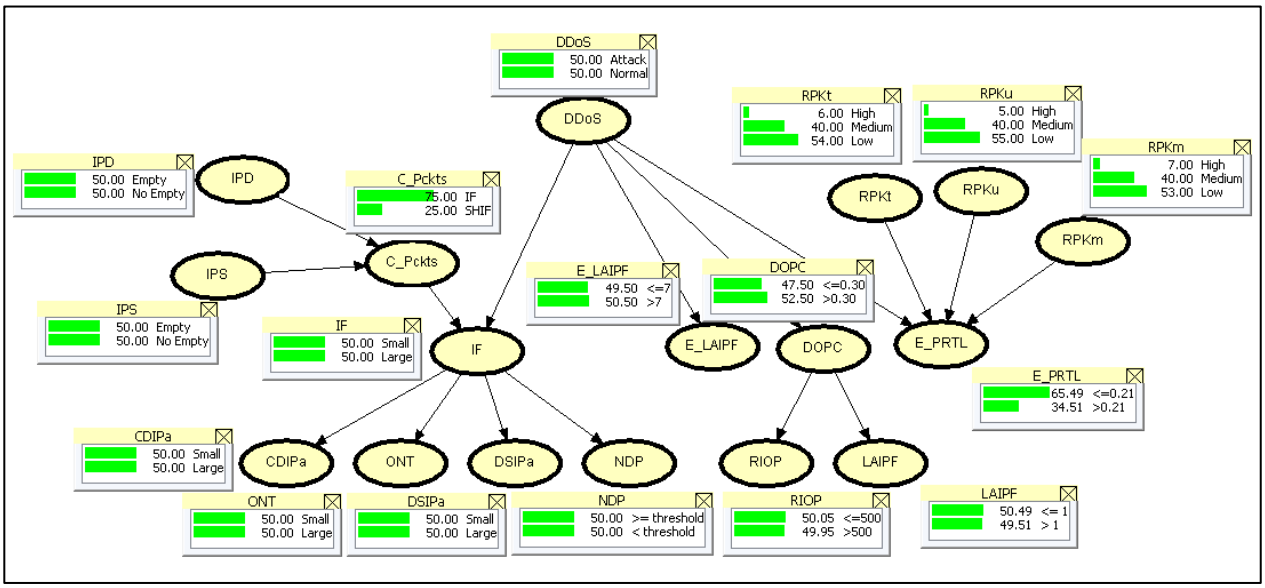

Figura 4.1.11 Modelo de la Red Bayesiana de la primera inferencia (P(DDoS/Attack)=0.5).

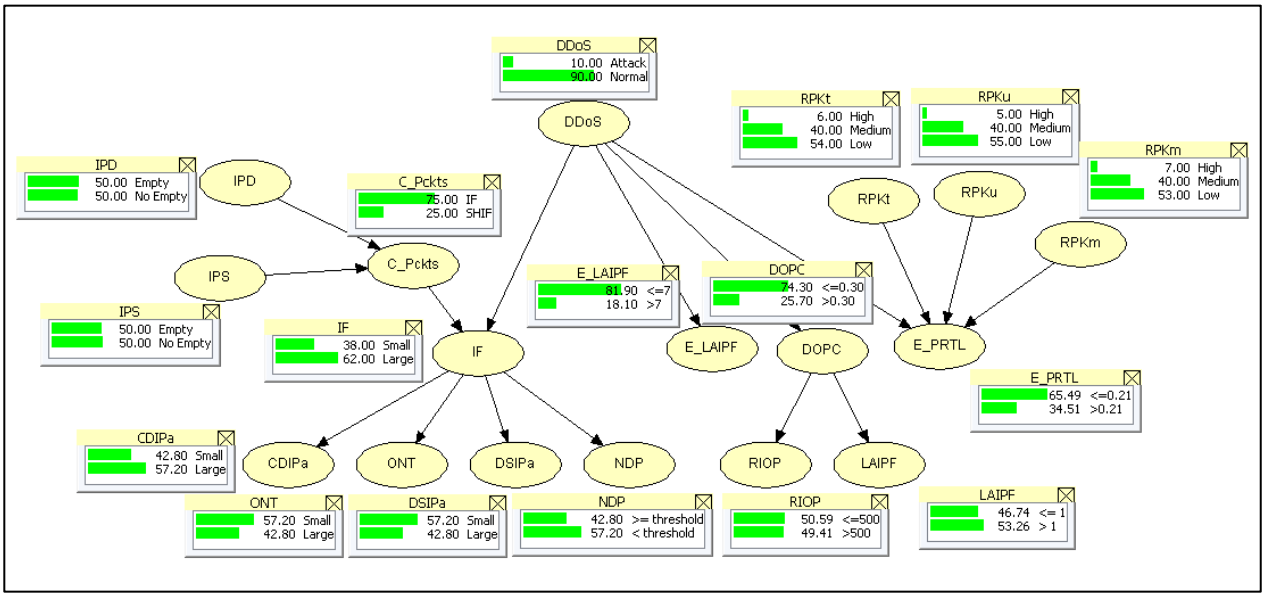

Figura 4.1.12 Modelo de la Red Bayesiana de la segunda inferencia (P(DDoS/Attack)=0.1).

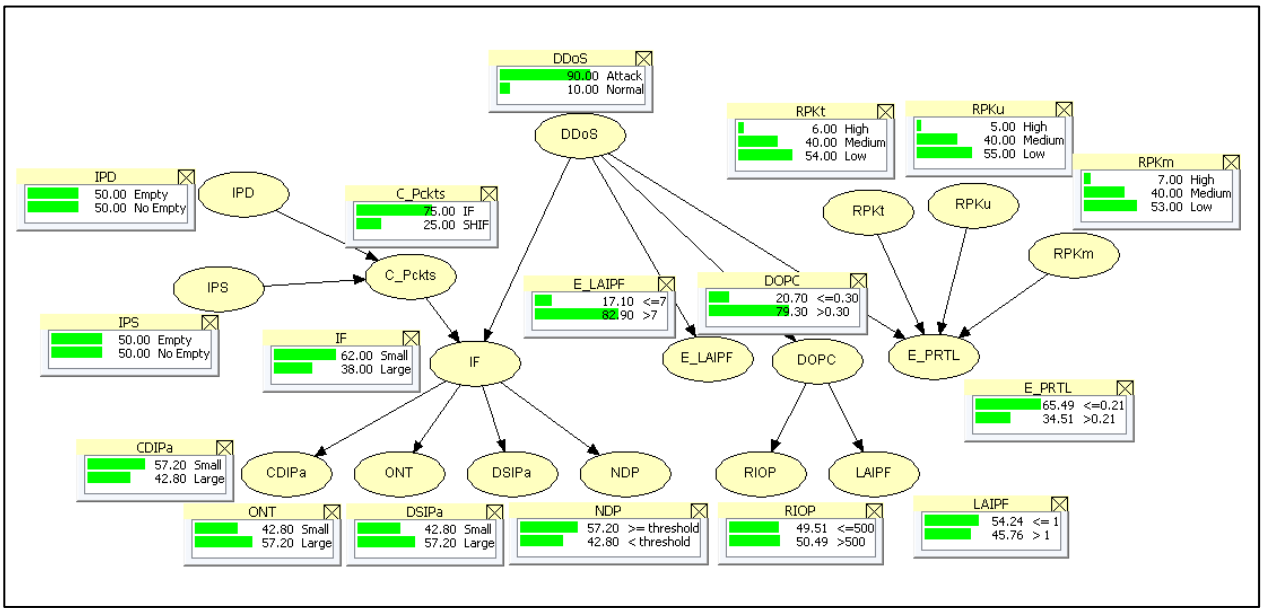

Figura 4.1.13 Modelo de la Red Bayesiana de la tercera inferencia (P(DDoS/Attack)=0.9).

Se eligieron tres escenarios posibles para el análisis inferencial de la variable aleatoria DDoS, los cuales se muestran en las Figuras 4.1.11 a la 4.1.13.

Como se puede observar cada uno de los escenarios ofrece una aproximación inferencial inicial distinta en algunos de sus nodos, originados por las características de las dependencias bayesianas entre nodos y el cálculo de auto propagación a través de la optimización triangular Naive Bayes.

Basados en el tráfico de paquetes que alimenta a la red bayesiana, se realizó una simulación de un ataque de DDoS, donde se generó evidencia que fue ingresada en las variables aleatorias de la red bayesiana para observar su comportamiento y proceso inferencial.

A continuación se muestran las redes bayesianas resultantes dada la evidencia de ataque por DDoS para los tres escenarios definidos:

I. Escenario 1: P(DDoS/Attack = 0.5

II. Escenario 2: P(DDoS/Attack = 0.1 y

III. Escenario 3: P(DDoS/Attack = 0.9

I. Para el Escenario 1

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia A*", que manifiesta un anomalía en el comportamiento de la interacción de las direcciones IP's del tráfico de datos:

a. IF es muy pequeño

b. NDP es mucho mayor al valor del umbral y

c. DSIPa es muy largo

La inferencia de la red bayesiana obtenida fue:

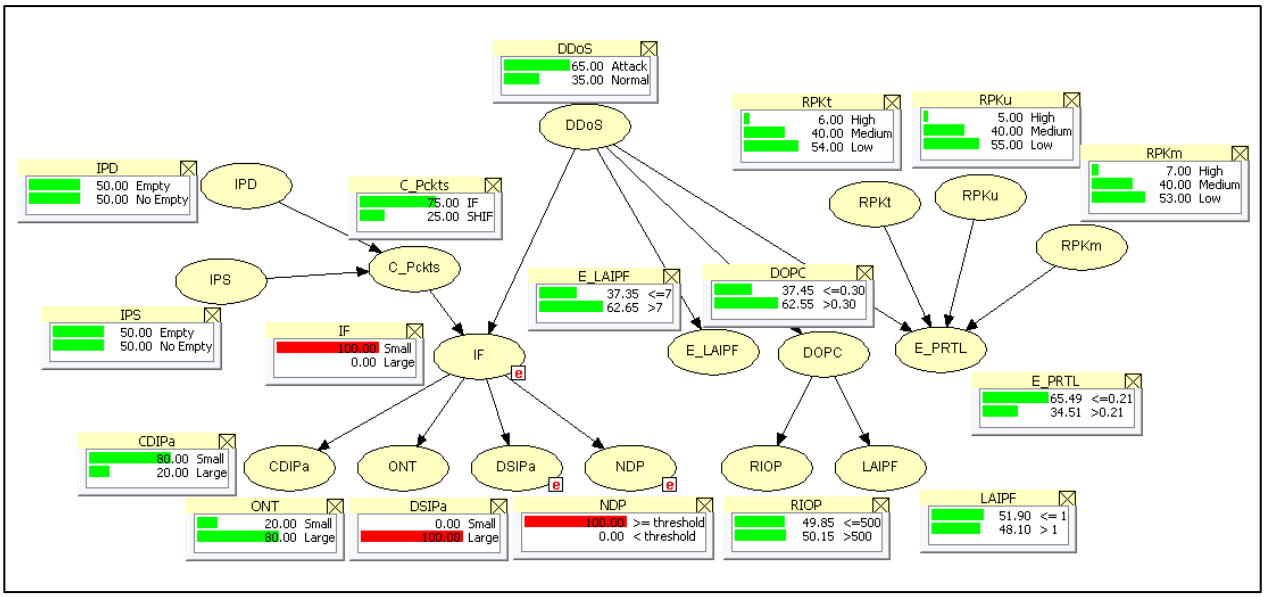

Figura 4.1.14 Modelo de la Red Bayesiana bajo el Escenario I y la Evidencia A.

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia B*", que manifiesta una anomalía en el comportamiento de la densidad de conexión de un solo flujo del tráfico de datos:

d. RIOP mayor a 500

e. LAIPF mayor a 1 y

f. E_LAIPF mayor a 7

La inferencia de la red bayesiana obtenida fue:

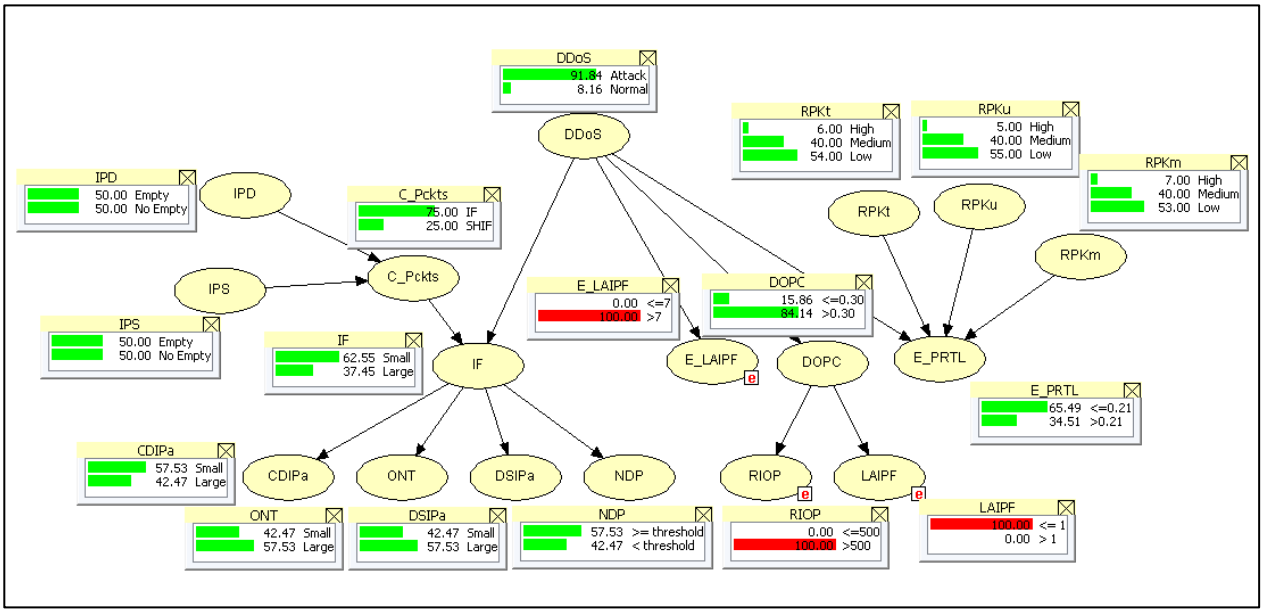

Figura 4.1.15 Modelo de la Red Bayesiana bajo el Escenario I y la Evidencia B.

II. Para el Escenario 2

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia A*", que manifiesta un anomalía en el comportamiento de la interacción de las direcciones IP's del tráfico de datos:

a. IF es muy pequeño

b. NDP es mucho mayor al valor del umbral y

c. DSIPa es muy largo

La inferencia de la red bayesiana obtenida fue:

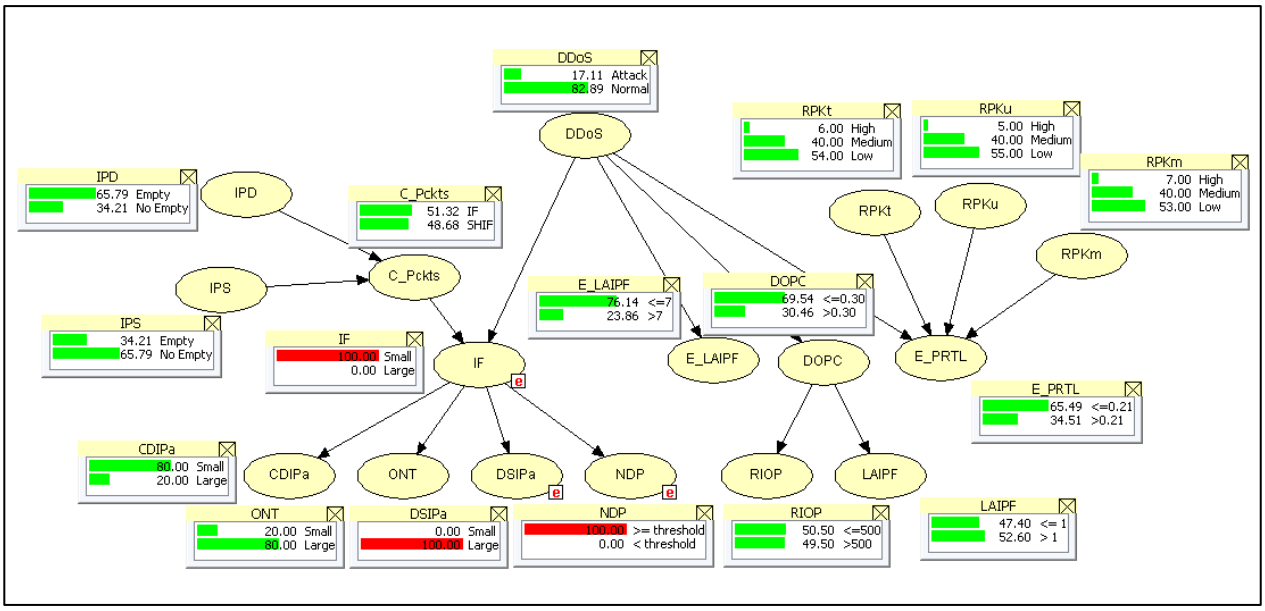

Figura 4.1.16 Modelo de la Red Bayesiana bajo el Escenario II y la Evidencia A.

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia B*", que manifiesta una anomalía en el comportamiento de la densidad de conexión de un solo flujo del tráfico de datos:

d. RIOP mayor a 500

e. LAIPF mayor a 1 y

f. E_LAIPF mayor a 7

La inferencia de la red bayesiana obtenida fue:

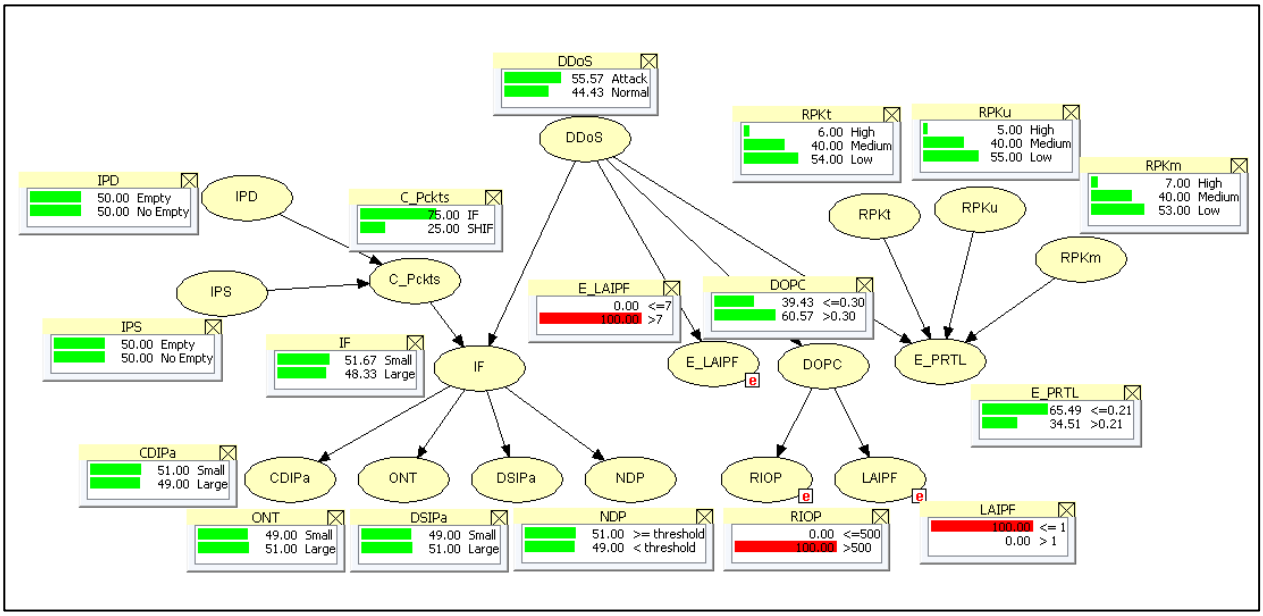

Figura 4.1.17 Modelo de la Red Bayesiana bajo el Escenario II y la Evidencia B.

III. Para el Escenario 3

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia A*", que manifiesta un anomalía en el comportamiento de la interacción de las direcciones IP's del tráfico de datos:

a. IF es muy pequeño

b. NDP es mucho mayor al valor del umbral y

c. DSIPa es muy largo

La inferencia de la red bayesiana obtenida fue:

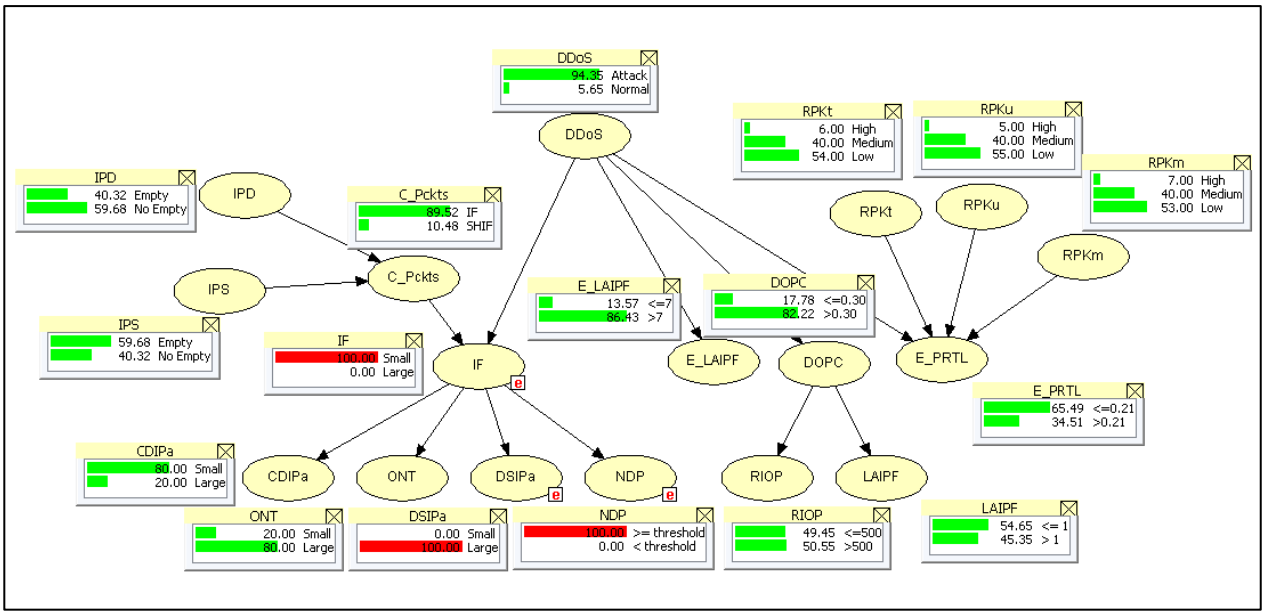

Figura 4.1.18 Modelo de la Red Bayesiana bajo el Escenario III y la Evidencia A.

Teniendo el siguiente conjunto de evidencia que llamaremos "*Evidencia B*", que manifiesta una anomalía en el comportamiento de la densidad de conexión de un solo flujo del tráfico de datos:

d. RIOP mayor a 500

e. LAIPF mayor a 1 y

f. E_LAIPF mayor a 7

La inferencia de la red bayesiana obtenida fue:

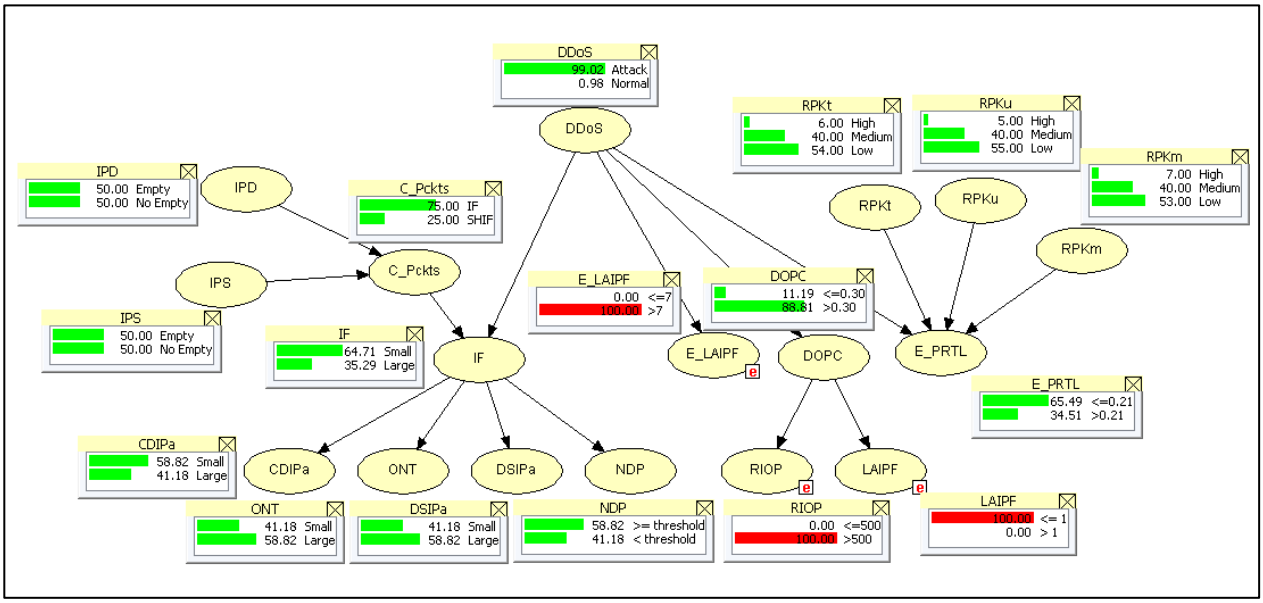

Figura 4.1.19 Modelo de la Red Bayesiana bajo el Escenario III y la Evidencia B.

Es posible continuar conformando conjuntos de evidencias para observar el comportamiento de la red bayesiana bajo estos parámetros, sin embargo, se consideró como suficiente la construcción de los escenarios I, II y III los cuales por ser valores medio y extremos en ambos estados de la variable aleatoria DDoS, dan un valor de significancia suficiente.

En cuanto a los grupos de evidencia A y B, fueron considerados como representativos basados en los experimentos realizados por Cheng, Yin, Cai & Wu, 2009 y de Xu, He & Lou, 2007, respectivamente, donde en sus trabajos hicieron aportaciones sobre los umbrales y comportamientos típicos de ataques de DDoS, para redes de datos con topologías similares a la que se propuso en esta investigación con las variables aleatorias seleccionadas para recolección de la evidencia.

Se realizó la propagación de inferencia de la red bayesiana considerando el escenario I y III, y la alimentación del conjunto de evidencias A y B, es decir, aplicando de forma integral ambos modelos de diagnóstico, se obtuvieron las gráficas de las redes bayesianas resultantes que se muestran en las Figuras 4.1.20 y 4.1.21

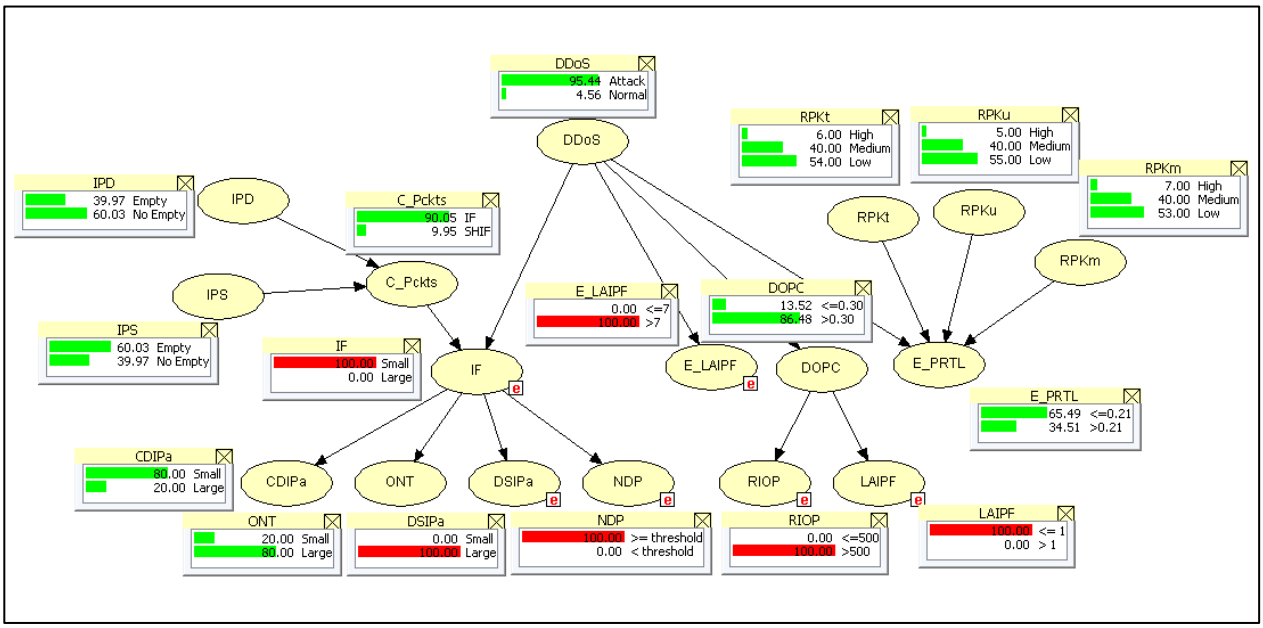

Figura 4.1.20 Modelo de la Red Bayesiana bajo el Escenario I y la Evidencia A y B.

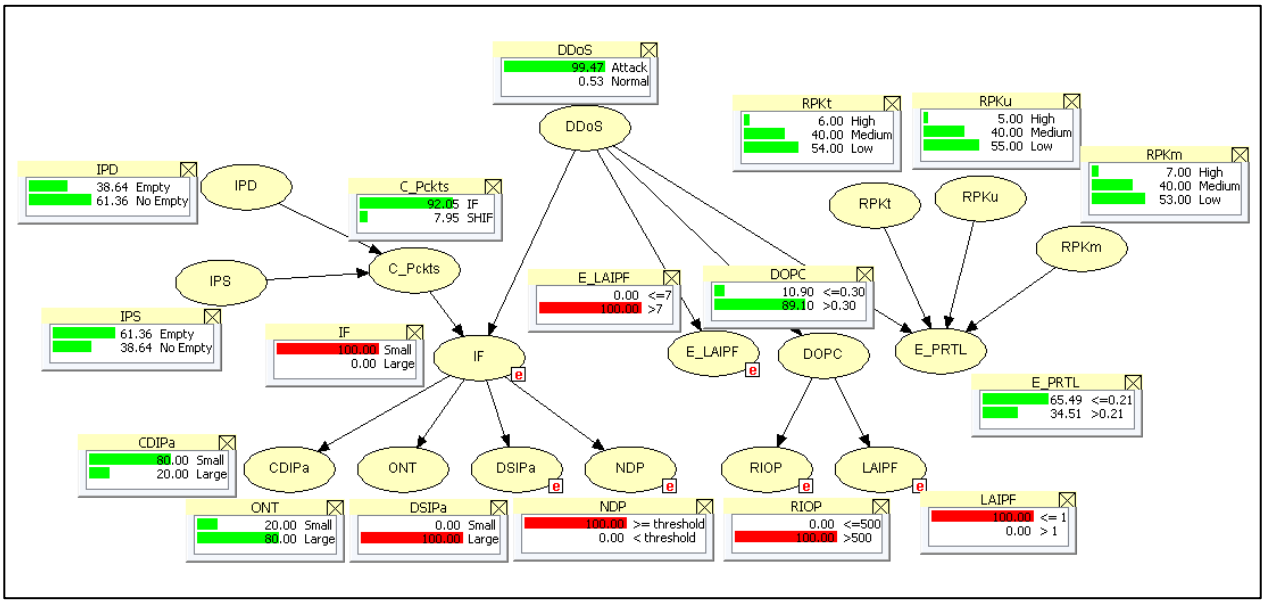

Figura 4.1.21 Modelo de la Red Bayesiana bajo el Escenario III y la Evidencia A y B.

4.3.3. Análisis de los resultados de la red bayesiana

Al construir una red bayesiana basado en dos modelos de diagnóstico de ataques de DDoS, cada uno de ellos con un conjunto específico de variables aleatorias, así como la aplicación de los mismos escenarios y evidencias para ambos, nos permitió observar el comportamiento individual de cada modelo para identificar un ataque de DDoS, y con ello poder ofrecer una opinión sobre su desempeño para inferir un ataque.

Como podemos observar en las gráficas de las Figuras 4.1.14, 4.1.16 y 4.1.18, la determinación *a priori* del valor sobre la variable de respuesta causal DDoS tiene una alta influencia, sobre las inferencias resultantes cuando se alimenta a la red bayesiana con la evidencia obtenida durante el estímulo de un ataque.

Cuando se consideran valores *a priori* bajos para el estado de "*ataque*" de la variable aleatoria DDoS, la red bayesiana requiere de mucha más evidencia positiva hacia la interpretación de un ataque, por lo que le costará mucho más tiempo, alcanzar un aprendizaje que le permita reducir la tasa de falsos negativos sobre la identificación de ataques por DDoS.

En cambio, cuando consideran valores *a priori* altos para el estado de "*ataque*" de la variable aleatoria DDoS, la red bayesiana requiere de mucha menos evidencia positiva hacia la interpretación de un ataque, por lo que le costará mucho más tiempo, alcanzar un aprendizaje que le permita reducir la tasa de falsos positivos sobre la identificación de ataques por DDoS.

Esto mismo ocurre para cualquier tipo de modelo de diagnóstico que se haya estimulado con la evidencia, se puede observar en las gráficas de las Figuras 4.1.15, 4.1.17 y 4.1.19 obtenidas a través del conjunto de evidencia B respecto a la red bayesiana resultante del conjunto de evidencia A.

También se puede observar que tiene una sensibilidad mayor para la estimación de probabilidades de la variable de respuesta causal DDoS, el modelo de diagnóstico utilizado para la creación del conjunto de evidencias B que el modelo de diagnóstico del conjunto de evidencias A.

Lo cual, no necesariamente nos puede llevar a pensar que un modelo es mejor que otro, sin embargo, nos permite presumir, que la combinación de ambos, puede ofrecer un grado de sensibilidad mucho más fino y preciso que de manera individual.

Lo anterior, nos permitirá alcanzar una convergencia hacia parámetros óptimos de inferencia para la detección de un ataque por DDoS más rápido y con menos entrenamiento y alcanzar tasas de falsos positivos mucho más bajas que, haciendo inferencias para cada modelo de forma individual.

También es posible determinar, que combinación de variables tienen mayor injerencia en el modelo integral de diagnóstico, permitiendo llegar a establecer el punto óptimo de sensibilidad para la identificación de un ataque de DDoS con estos modelos utilizados en la red bayesiana.

Como se puede observar en las gráficas de las Figuras 4.1.20 y 4.1.21, los escenarios I y III, que fueron los que mejor desempeño ofrecieron en sus estimaciones de ataque, alimentándolos con las evidencias correspondientes a los dos modelos de diagnóstico de forma integral (evidencias A y B), ofrecen mejores inferencias probabilísticas de la identificación de un ataque de DDoS respecto a las inferencias de los modelos de forma individual.

CONCLUSIONES

Este trabajo ha permitido ofrecer una esquema diferente para la identificación de anomalías y ataques a una red de datos por el método de DDoS, a través del desarrollo de un modelo basado en redes bayesianas.

Las redes bayesianas si bien es cierto que han sido utilizadas en áreas de la ciencias forenses (Taroni, Aitken, Garbolino, & Biedermann, 2010; Taroni, Bozza, & Aitken, 2005; Taroni, Bozza, & Biedermann, 2005; Rekhis, Krichene, Boudriga, 2009; Pourret, Naïm, & Marcot, 2010), existe pocos trabajos que presenten su utilización para la identificación de un ataque a una red de datos.

La hipótesis sobre la posibilidad de modelar una red bayesiana para identificar un ataque de DDoS en una red de datos, ha podido ser confirmada a través del presente trabajo y de los resultados obtenidos, los cuales son mostrados en las gráficas de las Figuras 4.1.20 y 4.2.21, con un nivel de confianza del 95% respecto a la distribución de probabilidades obtenidas para cada variable aleatoria utilizada en la red bayesiana propuesta.

También podemos argumentar, que una red bayesiana puede ser una herramienta eficiente para la identificación de ataques on-line a una red de datos, puesto que requiere un procesamiento mínimo para ser alimentada y calculados sus parámetros de inferencia, lo que permite ofrecer una opción con un bajo nivel de *over head* al desempeño de la red completa.

Puesto que las redes bayesianas se sustentan en el principio de dependencia condicional de las probabilidades de cada una de sus variables aleatorias con las que fueron modeladas, nos permite una facilidad de utilización cuando se cuenta con un modelo de comportamiento que puede generar un aprendizaje, tal y como un atacante se puede comportar para intentar vulnerar una red de datos.

Las redes bayesianas permiten también trasladar el conocimiento previo de expertos y de la experiencia adquirida del comportamiento típico de cualquier tipo de red, ofreciendo un punto de partida para el análisis estadístico de una serie de variables de

interés, que permita construir un estado *a priori* de inferencia para identificar un posible comportamiento anómalo.

Las redes bayesianas pueden crecer incorporando variables aleatorias que permitan diagnosticar más evidencia, y mejorar las tasas de identificación de ataques, de falsos negativos y por supuesto de falsos positivos, que permitan generar confianza en la capacidad de identificación de anomalías e incorporar nuevos comportamientos de los atacantes.

También se pueden replicar los modelos de redes bayesianas a lo largo de cualquier topología de una red de datos, lo que permitirá recolectar una mayor cantidad de información para clasificarla y convertirla en evidencia que alimente a la red bayesiana y permita actualizar sus parámetros de inferencia y realizar cálculos de ajuste en su auto propagación.

Como cualquier modelo de aprendizaje, las redes bayesianas son una opción interesante para la identificación de ataques sobre activos tecnológicos a partir de un conjunto de evidencia empírica que puede servir no solo para el posterior análisis forense de las redes de datos, sino también para la evaluación del riesgo tecnológico que puede enfrentar una organización.

Es importante considerar que este trabajo permite ofrecerle al análisis forense, una mayor calidad y cantidad de información para su ejecución, que permita identificar el esquema de ataque y a la posible conformación de evidencia para su presentación ante las autoridades correspondientes.

RECOMENDACIONES

Recolectar información longitudinalmente de una red en producción de alguna organización con tráfico e intentos de ataque reales, para someter al modelo a una revisión y comparación con datos operativos reales.

Construir una base de datos institucional de tráfico real, para ser utilizada en el entrenamiento de modelos de identificación de anomalías y ataques basados en redes bayesianas, para que los trabajos futuros sobre esta línea de investigación tengan mayor certidumbre en sus hallazgos e inviertan menos tiempo en este proceso.

Incorporar a la red bayesiana propuesta, nuevos modelos de diagnóstico para comprobar el desempeño y la calidad de identificación de ataques de cualquier tipo a una red de datos.

Construir un modelo de identificación de ataques de DDoS a una red de datos basado en redes neuronales para ofrecer un comparativo con el modelo propuesto en las características relativas al desempeño, calidad, eficiencia y efectividad en la identificación de ataques por DDoS.

Continuar con el diseño y construcción de una HoneyNet que pueda ser receptora del tráfico y de los equipos identificados como posibles atacantes, para continuar la recolección de evidencia que permita realizar un análisis forense completo y amplio.

Continuar con el diseño y construcción de la interface de bajo nivel para la inicialización de los dispositivos de grabación de bitácoras que recolecten la evidencia del posible ataque identificado por la red bayesiana.

Construir un modelo de asignación de las ubicaciones físicas en una topología de red donde deben instalarse los modelos de identificación de ataques de DDoS basado en redes bayesianas que este trabajo propone.

LIMITACIONES Y RETOS

Este trabajo solamente consideró un solo tipo de ataque a una infraestructura de red conocido como DDoS, a través de dos modelos de diagnóstico utilizados en los últimos 3 años.

No fue posible construir una base de datos con información mucho más amplia longitudinalmente y en cantidad, para considerar un espectro más amplio de posibilidades de realización de ataques de DDoS que añadiera mayor confiabilidad al modelo propuesto.

La base de datos que sirvió para la construcción de los parámetros *a priori* del modelo de la red bayesiana, fue completamente experimental, por lo que las inferencias obtenidas y presentadas en el presente trabajo, deben ser calculadas nuevamente si se utiliza este modelo en una red de datos en producción con tráficos reales de operación.

Uno de los retos que este trabajo enfrentó, fue la obtención de los permisos para la utilización de una infraestructura tecnológica, para la creación de la red experimental y la construcción de la base de datos del tráfico, que fue necesaria para la realización del análisis estadístico que generó los valores *a priori* de la red bayesiana para su auto propagación inferencial.

Otro de los retos importantes que este trabajo tuvo fue, el diseño y construcción de la red bayesiana que incorporó dos modelos de diagnóstico de ataques de DDoS probados con anterioridad, para la identificación de ataques de este tipo.

REFERENCIAS

1. Achi, H., Hellany, A. & Nagrial, M. (2008). Network Security Approach for Digital Forensics Analysis. *Technical Report IEEE Computer Society 978-1-4244-2116-9/08*.
2. Al-Duwairi, B. & Daniels, T. (2004). Topology-based Packet Marking, ICCCN, 2004.
3. Amran, A.R., Phan, R.C.W. & Parish, D.J. (2009). *Metrics for Network Forensics Conviction Evidence. Technical Report*. Department of Electronic & Electrical Engineering, Loughborough University UK, 2009. IEEE Computer Society
4. Aitken, C., Gammerman, A., Zhang, G., Connolly, T., Bailey, D., Gordon, R. & Oldfield, R. (1996). Bayesian belief networks with an application in specific case analysis *In Computational Learning and Probabilistic Reasoning* (ed. Gammerman, A.), pp 169-184. John Wiley & Son, Chichester.
5. Arun Raj Kumar, P. & Selvakumar, S. (2009). Distributed denial of service (DDoS) threat in collaborative environment a survey on DDoS attack tools and traceback mechanisms. *2009 IEEE International Advance Computing Conference (IACC 2009)*. India March 2009.
6. Ashif, A., Alam, O.M. & Aktaruzzaman, A.K.M. (2009). TCP SYN Flood DoS Attack Experiments in Wireless Network. *IEEE International Advance Computing Conference (IACC 2009)*. India March 2009.
7. Balkema, A.A. & de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2:792-804, 1974. Retrieved March 23, 2011, from http://projecteuclid.org/euclid.aop/1176996548
8. Belenky, A. & Ansari, N. (2003) IP Traceback with Deterministic Packet Marking, *IEEE Communication Letters*, vol 7(4), 2003.
9. Belloving, M., S. (2000). ICMP Traceback Messages, Internet Draft: draft-bellovinitrace-00.txt, 2000.
10. Burch, H. & Cheswick, B. (2000). Tracing Anonymous Packets to their approximate source, *USENIX LISA, 2000*.
11. Celenk, M., Conley, T., Willis, J., Graham, J. (2010). Predictive Network Anomaly Detection and Visualization. *IEEE Transactions on Information Forensics and Security*, Vol. 5, No. 2, June 2010. IEEE Computer Society 1556-6013.
12. Cheng, J., Yin, J., Liu, Y., Cai, Z., Wu, C. (2009). DDoS attack detection using IP address feature interaction. *2009 International Conference on Intelligence Networking and Collaborative Systems*. IEEE Computer Society 978-0-7695-3858-7/09.
13. Cho, K., Kaizaki, R., & Kato, A. (2002). An aggregation technique for traffic monitoring. *Symp. Applications and the Internet (SAINT) Work shops*, 2002, p. 74.
14. Cornalba, C. & Giudici, P. (2004). Statistical Models for Operational Risk Management, Physical A: Statistical Mechanics and its applications. Retrieved March 12, 2011, from

www.consorziocarma.com/pia/index.php?option=com_docman&task=doc_view&gid=11&Itemid=12

15. Correa, M., Bielsa, C., Teixeira-Pamies, J., & Alique, J.R. (2006). Redes Bayesianas vs Redes Neuronales en Modelos para la Predicción del Acabado Superficial. *Technical Note IEEE Automation Society 3244-5445*.

16. Darwiche, A. (2009). *Modeling with Bayesian Networks*. Cambridge University Press; 1st. Edition, New York, NY.

17. Dean, D. et. al. (2000). An Algebraic Approach to IP Traceback, *ACM TISSEC*, 5(2), pp. 119-137, 2000.

18. Eimann, R., Speidel, U., Brownlee, N. & Yang, J. (2005). Network Event Detection With T-Entropy. *Centre for Discrete Mathematics and Theoretical Computer Science*, University of Auckland, New Zealand, Rep. CDMTCS-266, May 2005.

19. El-Shehaly, M., Gracanin, D., Abdel-Hamid, A., Kresimir, M. (2009). A visualization framework for Traffic Data Exploration and Scan Detection. *IEEE Computer Society 978-1-4244-6273-5/09*.

20. Feldmann, A., Gilbert, A. C. & Willinger, W. (1998). Data networks as cas cades: Investigating the multifractal nature of internet WAN traffic. *in Proc. SIGCOMM*, 1998, pp. 42–55.

21. Fitzmaurice, G. M., Davidian, M., Verbeke, G. & Molenberghs, G. (2008). *Longitudinal Data Analysis*. Chapman & Hall/CRC, Taylor & Francis Group. Boca Raton, FL.

22. Foresti, S., Agutter, J., Livnat, Y., Moon, S., & Erbacher, R.F. (2006). Visual correlation of network alerts. *IEEE Computing Graphics Applications*, Vol. 26, No. 2, pp.48-59. Apr. 2006.

23. Friedman, N., Geiger, D., & Goldszmidt (1998). Bayesian Network Classifiers. Kluwer Academic Publisher, Boston. 1997, 1-37(). Retrieved March, 11 2011 from: http://www.cs.huji.ac.il/~nir/Papers/FrGG1.pdf

24. Gianvecchio, S. & Wang, H. (2007). Detecting covert timing channels: An entropy-based approach. *Proc. ACM Conf. Computer and Communications Security*, 2007, pp. 307–316.

25. Gong, Y. (2004A). Security Focus Article: Detecting Worms and Abnormal Activities With NetFlows, Part 1 Aug. 2004 [Online]. Available: http:// www.securityfocus.com/infocus/1796.

26. Gong, Y. (2004B). Security Focus Article: Detecting Worms and Abnormal Ac tivities With NetFlows, Part 2 Sep. 2004 [Online]. Available: http:// www.securityfocus.com/infocus/1796.

27. Gu, Y., McCallum, A. & Towsley, D. (2005). Detecting anomalies in net work traffic using maximum entropy estimation. *Proc. 5th ACM SIGCOMM Conf. Internet Measurement (IMC '05)*, New York, 2005, pp. 1–6, ACM.

28. Hajji, H. (2005). Statistical analysis of network traffic for adaptive faults de tection. *IEEE Trans. Neural Netw.*, vol. 16, no. 5, pp. 1053–1063, Sep 2005.

29. Harrington, E.F. (2006). Measuring network change: Rényi cross entropy and the second order degree distribution. *Proc. Passive and Active Measurement (PAM) Conf.*, Adelaide, Australia, Mar. 2006.

30. Hedeker, D & Gibbons, R. (2006). *Longitudinal Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey.

31. Heckermand, D. (1995). A tutorial on learning with Bayesian Networks. *Technical Report MSR-TR-95-06*, Microsoft Research, Advanced Technology Division, March 1995.

32. Hernández, R., Fernández, C., & Baptista, P. (2002). *Metodología de la Investigación.* México: 4ª. Edición, Ed. McGraw-Hill.

33. ICCC Internet Crime Complaint Center (2010). 2010 Internet Crime Report. Retrieved March 25, 2011 from: http://ic3report.nw3c.org/

34. Jin, S., Yeung, D. S. & Wang, X. (2007). Network intrusion detection in covariance feature space. *Pattern Recognit.*, vol. 40, no. 8, pp. 2185–2197, 2007.

35. Jensen, F. V. Nielsen T.D. (2001). *Bayesian Networks and decision graphs*. Springer-Verlag, New York, Inc. March 2007.

36. Kan, S., Kim, J. (2008). Network Forensic Analysis Using Visualization Effect. *International Conference on Convergence and Hybrid Information Technology 2008*. IEEE Computer Society 978-0-7695-3328-5/08, p466-473.

37. Karasaridis, A., Rexroad, B. & Hoeflin, D. (2007). Wide-scale botnet detec tion and characterization. *Proc. First Conf. First Workshop on Hot Topics in Understanding Botnets (HotBots'07)*, Berkeley, CA, 2007, p. 7, USENIX Association.

38. Kim, S. S., Reddy, A. L. N. & Vannucci, M. (2004). Detecting traffic anomalies through aggregate analysis of packet header data. *Networking*. New York: Springer, 2004, vol. 3042, pp. 1047–1059.

39. Kim, S. S. & Reddy, A. L. N. (2008). Statistical techniques for detecting traffic anomalies through packet header data. *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 562–575, Jun. 2008.

40. Kjaerulff, B. U., & Madsen, A. (2008). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Springer Science+Business Media LLC, New York, NY.

41. Kwit, R., Hofmann, U. (2007). Unsupervised anomaly detection in network traffic by means of robust PCA. *In Proc. Int. Conf. Computing in the Global Information Technology (ICCGI)*, 2007, pp. 37-39.

42. Lakhina, A., Crovella, M. & Diot, C. (2005). Mining Anomalies Using Traffic Distributions CS Department, Boston University, Tech. Rep. 2005-002, Feb. 2005.

43. Lall, A., Sekar, V., Ogihara, M., Xu, J. J. & Zhang, H. (2005). Data Streaming Algorithms for Estimating Entropy of Network Traffic Computer Sci ence Department, University of Rochester, Tech. Rep. TR886, Nov. 2005.

44. Lin, C., Zhintang, L., Cuixia, G. (2009). Automated Analysis of Multi-source Logs for Network Forensics. *2009 First International Workshop on Education Technology and Computer Science*. IEEE Computer Society 978-0-7695-3557-5/09, 660-664.

45. Margaritis, D. (2003). Learning Bayesian Network Model Structure from Data. *Technical Report CMU-CS-03-153*, Carnegie Mellon University, School of Computer Science, May 2003.

46. McNeil, A. (1997). Estimating the tails of loss severity distributions using extreme value theory, *ASTIN Bulletin* **27**, 117-137. Retrieved March 14, 2011 from www.ma.hw.ac.uk/~mcneil/ftp/astin.pdf

47. McNeil, A. (1999). Extreme Value Theory for Risk Managers, ETH Zentrum Zürich. Retrieved March 14, 2011 from www.math.ethz.ch/~mcneil/ftp/cad.pdf

48. Ming, H., LiZhong, S. (2009). A New System Design of Network Invasion Forensics. *2009 Second International Conference on Computer and Electrical Engineering*. IEEE Computer Society 978-0-7695-3925-6/09, 596-599

49. Muthuprasanna, M. & Manimaran, G. (2005). Space-Time Encoding Scheme for DDoS Attack Traceback*, IEEE GLOBECOM, 2005*.

50. Muthuprasanna, M. & Manimaran, G. (2008). Distributive divide and conquer techniques for effective DDoS attack defenses*, The 28th International Conference on Distributed Computing Systems. IEEE GLOBECOM, 2008*.

51. Nehinbe, J.O., (2010). Log Analyzer for Network Forensics and Incident Reporting. *2010 International Conference on Intelligent Systems, Modeling and Simulation*. IEEE Computer Society 978-0-7695-3973-7/10, p356-361.

52. Nelson, W (1990). *Accelerated Testing, Statistical Models: Test Plans and Data Analyses*. United States, New York, Ed. Wiley Interscience.

53. Pang, R., Yegneswaran, V., Barford, P., Paxson, V., & Peterson, L. (2004) Characteristics of internet background radiation. *Proc. ACM In ternet Measurement Conf.*, Oct. 2004, pp. 27–40.

54. Park, D.W. (2008). A Study of packet analysis regarding a DDoS attack in WiBro environments. *IJCSNS International Journal of Computer Science and Network Security*, Vol 8., No.12, December 2008.

55. Plonka, D. (2000). A network traffic flow reporting and visualization tool. *Proc. 14th USENIX Conf. System Administration*, New Orleans, LA, 2000, pp. 305–318.

56. Pourret, O., Naïm, P. & Marcot, B. (2010). *Bayesian Networks: A practical Guide to Applications*. John Wiley & Sons, Ltd. The atrium, Southern Gate, Chichester West Sussex PO19 85Q, England.

57. Rebane, G. & Pearl, J. (1987). The recovery of causal poly-trees form statistical data. Technical Report UCLA. Retrieved March 12, 2011, from: *ftp.cs.ucla.edu/tech-report/198_-reports/870031.pdf*

58. Rekhis, S., Krichene, J., Boudriga, N. (2009). Forensic Investigation in Communication Networks Using Incomplete Digital Evidences. *International Journal Communication, Network and System Science*, 2009, 2, 857-873

59. Sang, A. & Li, S. (2000) A predictability analysis of network traffic. *Proc. INFOCOM (1)*, 2000, pp. 342–351.

60. Savage, S., Wetherall, D., Karlin, A. & Anderson, T. (2000). Practical Network Support for IP Traceback, *ACM SIGCOMM, 2000*.

61. Shen, G., Chen, D., & Qin, Z. (2007). Anomaly detection based on aggregated network behavior metrics. *Proc. Wireless Communications, Net working and Mobile Computing, 2007 (WiCom 2007)*, pp. 2210–2213.

62. Snoeren, A., et. al. (2002). Single-Packet IP Traceback, *IEEE/ACM Trans. On Networking*, 10(6), pp. 721-734, 2002.

63. Song, D. & Perrig, A. (2001). Advanced and Authenticated Marking Schemes for IP Traceback, *IEEE INFOCOM*, 2001.

64. Stone, R. (2000). Center Track: An IP overlay network for tracking DDoS floods, *USENIX Security Symposium, 2000*.

65. Taroni, F., Aitken, C., Garbolino, P., Biedermann, A. (2010). *Bayesian Networks and probabilistic inference in Forensic Science*. John Wiley & Sons, Ltd. The atrium, Southern Gate, Chichester West Sussex PO19 85Q, England.

66. Taroni, F., Bozza, S., & Aitken, C. (2005). Decision analysis in forensic science. *Journal of Forensic Sciences* **50**, 894-905.

67. Taroni, F., Bozza, S., & Biedermann, A. (2005). Two items of evidence, no putative source: an inference problem in forensic intelligence. *Journal of Forensic Science* **55**, 456-467.

68. Thottan, M. & Ji, C. (2003). Anomaly detection in IP networks. *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2191–2204, Aug. 2003.

69. Verbeke, G. & Molenberghs, G. (2009). *Linear Mixed Models for Longitudinal Data*. Springer Verlag, New York, LLC.

70. Wagner, A. & Plattner, B. (2005). Entropy based worm and anomaly de tection in fast IP networks. *Proc. 14th IEEE Int. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise*, Washington, DC, 2005, pp. 172–177.

71. Xu, T., He, D., Luo, Y. (2007). DDoS Attack detection based on RLT features. *2007 International Conference on Computational Intelligence and Security*. IEEE Computer Society 0-7695-3072-9/07, p697-701.

72. Xu, T., He, D. Zheng, Y. (2006). Detecting DDoS attack based on one-way connection density. *IEEE Computer Society* 1-4244-0411-8/06.

73. Yasuda, Y. (2003). *Application of Bayesian Inference to Operational Risk Management*. Master Thesis for Doctoral Program in Quantitative Finance and Management, for University of Tsukuba.

74. Yurcik, W. & Li, Y. (2005). Internet security visualization case study: Instru menting a network for netflow security visualization tolos. *Proc. Annual Computer Security Applications Conf. (ACSAC 05)*, Tucson, AZ, Dec. 5–9, 2005.

75. Zhang, Y., Liu, Q., Zhao, G. (2010). A real-time DDoS attack detection and prevention system based on per-IP traffic behavioral analysis. IEEE Computer Society 978-1-4244-5540-9/10, p163-167.

76. Zhang, R. & Bivens, A. (2007) Comparing the use of Bayesian networks and neural networks in respond time modeling for service-oriented systems. *Proceedings of the 2007 Workshop on Service oriented Computing Performance: Aspects, Issues and Approaches*. 2007: Monterey. p. 67-74.

77. Ziviani, A., Monsores, M.L., Rodrigues, P.S.S., Gomes, A.T.A (2007). Networks Anomaly detection using nonextensive entropy, IEEE Commun. Lett., vol. 11, no. 12, pp.1034-1036, Dec. 2007.