Capítulo 1. Introducción

En este capítulo se presenta una descripción general del problema a investigar y el enfoque con el que se aborda. Se establece la necesidad de incorporar técnicas de análisis novedosas a los sistemas de detección de intrusos que funcionan en una red de comunicación y se da una breve reseña de los trabajos previos en esta área. En la parte final del capítulo se describen los objetivos del proyecto y la organización de los capítulos que integran este documento.

1.1. Antecedentes

El acceso creciente a las redes globales de comunicación como Internet ha permitido que los usuarios accedan a numerosas fuentes de información alrededor del mundo y mantengan contacto con distintas personas de manera rápida, sin embargo como señalan las estadísticas presentadas por el equipo especializado en seguridad informática C.E.R.T. [CER10], en los últimos años ha existido un incremento en el número de ataques o conductas intrusivas que han logrado infringir las tecnologías de protección establecidas en distintas organizaciones. Debido a esto, ha surgido la necesidad de crear nuevos mecanismos que permitan detectar ataques a la infraestructura de red y así proteger la información electrónica de organizaciones y usuarios.

Los sistemas de seguridad consideran distintos niveles de protección. Los cortafuegos o firewalls son dispositivos que controlan el acceso entre distintos segmentos de red a través de una serie de reglas establecidas. Otro esquema de protección considera la definición de zonas o segmentos de red donde se concentren los dispositivos más vulnerables a recibir ataques como servidores web o proveedores de servicios, a esta zona se le conoce con el nombre de zona desmilitarizada (DMZ).

Los Firewalls, zonas desmilitarizadas y listas de control de acceso forman la primera línea de defensa en las redes de computadoras, sin embargo estos recursos pueden sufrir vulnerabilidades por lo que es necesario realizar investigación para desarrollar otros niveles de protección. Los sistemas de detección de intrusos mantienen un monitoreo constante sobre el tráfico de la red, ya sea que estén establecidos dentro de una red de área local (LAN) o entre distintos dominios de colisión. Generalmente tienen una entidad que administra la toma de decisiones en caso de la detección de conductas intrusivas. Como tarea principal, los sistemas de detección de intrusos brindan protección en tiempo real contra distintos tipos de ataques ya sean internos o externos. Además, pueden tener la

capacidad de ser reactivos a los ataques antes de que puedan ocasionar daños críticos en la red. En la Figura 1.1 se muestran las posibles configuraciones de un sistema de detección de intrusos (IDS) dentro de una topología de red.

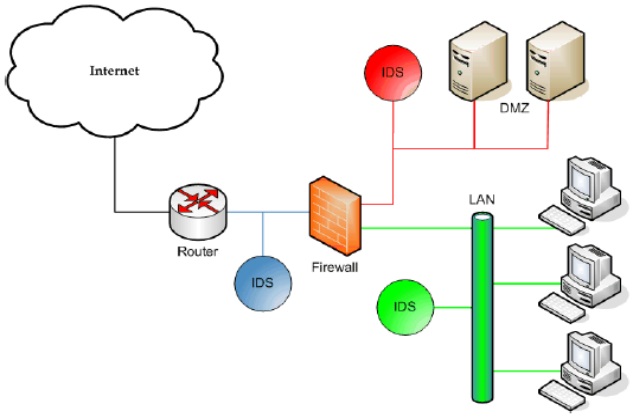

Figura 1.1: Topología de red.

En los últimos años la investigación en los sistemas de detección de intrusos se ha inclinado hacia el área de inteligencia artificial, dentro de éste campo se aprovechan las características de adaptabilidad y aprendizaje que tienen las redes neuronales para implementar tecnologías de detección de intrusos más dinámicas que mejoren el índice de detección de ataques o conductas intrusivas. Dentro de la técnica de detección de intrusos basada en anomalías se consideran algoritmos no lineales como las redes neuronales [PLA01]. Las redes neuronales basadas en wavelets son un nuevo tipo de redes neuronales que utilizan la transformada wavelet, combinan las características de localización de la transformada wavelet con el enfoque no lineal de las funciones de redes neuronales. La transformada wavelet es usada para reemplazar la función de transferencia en las unidades de procesamiento.

1.2. Planteamiento del problema

Para mejorar el índice de detección de ataques y conductas intrusivas sobre redes de comunicación, es necesario explorar nuevos modelos que se puedan aplicar a IDS que utilizan redes neuronales en su funcionamiento. Dado que en trabajos previos los mejores resultados se han obtenido utilizando redes neuronales recurrentes resulta natural entonces que en este trabajo se haga un énfasis especial en la investigación de este tipo de arquitecturas.

El uso de redes neuronales junto con el procesamiento wavelet busca aprovechar las características de ambos conceptos, es decir, utilizar las características de reconocimiento

de patrones y clasificación de las redes neuronales junto con las propiedades de representación y localización que poseen las funciones wavelet. Al definir un nuevo tipo de unidad de procesamiento es necesario describir su funcionamiento, es decir, la forma en la que produce la salida a partir de un estimulo o entrada dado, además de indicar las restricciones en sus parámetros. Las unidades de procesamiento junto con otros elementos como unidades de entrada y salida, forman parte de un esquema de interconexiones que definen el mapeo entrada – salida descrito por medio de una ecuación.

Una vez definida la parte estructural del modelo es necesario definir los parámetros libres de la arquitectura y sus propiedades. Durante la etapa de aprendizaje estos parámetros ajustarán sus valores para minimizar el error producido por el modelo, dichos ajustes son descritos mediante un algoritmo de entrenamiento.

Ésta tesis tiene como propósito modelar y simular un sistema de detección de intrusos que use redes neuronales recurrentes con wavelets, obteniendo resultados cuantitativos que permitan llevar a cabo un análisis comparativo respecto a otros trabajos previamente realizados. Las redes neuronales basadas en wavelets reemplazan la función sigmoide utilizada por las unidades en las capas de ocultas por una función wavelet, dando lugar a un nuevo tipo de unidad de procesamiento que implementa sus funciones de transferencia a partir de una wavelet madre [SUN09].

Este trabajo pretende investigar la utilización de una red neuronal recurrente basada en wavelets para detectar conductas intrusivas o ataques sobre una red de comunicación. Una de las principales funciones del modelo será categorizar tráfico de red como dañino o inofensivo, adicionalmente, dentro de los patrones clasificados como intrusivos describir la clase de ataque a la que pertenece. Durante la investigación se explorará el desempeño de las redes neuronales basadas en wavelets como complemento a tecnologías de seguridad existentes como sistemas de autentificación, seguridad de ruteo y firewalls para mejorar la precisión de detección de nuevos ataques y de esta manera incrementar el nivel de seguridad en el dominio donde se implemente el sistema.

La principal característica del enfoque adoptado para la presente investigación es el cambio de la función de transferencia sigmoide que utilizan las redes neuronales tradicionales por una función wavelet, y de este modo determinar si el modelo de red neuronal propuesto obtiene mejoras en la velocidad de convergencia e índices estadísticos comparativos para los sistemas de detección de intrusos. La construcción del modelo planteado anteriormente implica una etapa de diseño donde se establezcan las características generales de la arquitectura así como la deducción del algoritmo de entrenamiento a usar, una etapa de implementación donde se llevará a cabo la programación mediante un lenguaje de alto nivel del diseño y finalmente una etapa de simulación donde se podrán obtener resultados cuantitativos sobre el desempeño de la red neuronal que permitan concluir la viabilidad de este modelo. El resultado final contemplará

un esquema de interconexiones con distintas capas que después de una etapa de entrenamiento tendrá la capacidad de reconocer amenazas dentro del segmento de red donde se ubica el sistema de detección de intrusos.

1.3. Trabajos Previos

A partir de los primeros trabajos en el área de detección de intrusos por Dening [DEN87], se han desarrollado distintos tipos de IDS, cada uno basado en mecanismos distintos. Como se indica en [BAI03], la investigación en ésta área contempla distintos enfoques como el estadístico, la generación predictiva de patrones y el uso de redes neuronales.

En el trabajo presentado en [BOT10], se implementa un IDS portable que hace uso de la base de datos de firmas de SNORT[1]. El desarrollo de un IDS que emplea un método estadístico para el pre procesado de datos junto con una red neuronal para la clasificación de fallas y ataques es reportado en [MAN02], dicho sistema opera automáticamente en distintas tecnologías de red.

Dentro del enfoque con redes neuronales, en [SAN05] se comparan tres modelos: el primero utiliza una red con alimentación hacia adelante entrenada con el algoritmo de resilient back propagation, el segundo modelo utiliza una red con alimentación hacia delante entrenada con el algoritmo de gradiente conjugado y finalmente se presenta un modelo con red Elman entrenada con resilient back propagation. Otro trabajo comparativo de IDS con redes neuronales se encuentra en [ALA08] donde se presenta una comparación entre tres distintos modelos de IDS que utilizan redes neuronales: el primero utiliza una red con alimentación hacia adelante entrenada con el algoritmo resilient back propagation, el segundo modelo utiliza una red recurrente simple (Elman) entrenada con el algoritmo Backpropagation y el último modelo implementa una arquitecura recurrente totalmente conectada entrenada con el algoritmo Real Time Recurrent Learning (RTRL). Destacando por sus resultados entre las demás la arquitectura recurrente totalmente conectada. En el trabajo de Efraín Torres [MEJ03] se reporta un sistema inmunológico[2] de detección de intrusos que funciona a nivel de protocolo HTTP, este IDS utiliza una red Elman como parte de su mecanismo para llevar a cabo la detección, identificación y clasificación de ataques. El uso de redes neuronales recurrentes en la extracción de características para formar un grupo de reglas que describan las características de patrones intrusivos es

[1] Snort es un sistema de detección de intrusos creado por Martin Roesch que monitorea el tráfico dentro de un segmento de red, ofrece características como análisis y filtrado de paquetes en tiempo real, monitoreo de puertos y análisis de protocolos.

[2] La naturaleza inmunológica describe un sistema de protección multinivel que basa su funcionamiento en el reconocimiento de patrones. Este reconocimiento no se efectúa por medio de una memoria estática sino por una memoria evolutiva y dinámica [sistema inmunológico].

planteada en [XUE04], la estructura de red neuronal utilizada en este trabajo es una modificación de la red recurrente Jordan.

1.4. Objetivos

Los objetivos del presente estudio son:

1.4.1. Objetivo General

Proponer un modelo de red neuronal recurrente con wavelets aplicado a la detección de intrusos en una red de comunicación de datos, dicho modelo contempla la definición de la arquitectura y el algoritmo de entrenamiento. El modelo planteado será capaz de analizar distintos patrones en la capa de aplicación del modelo OSI[3] y llevará a cabo las tareas de detección, identificación y clasificación de ataques.

1.4.2. Objetivos Específicos

En la actualidad, la mayoría de los IDS basan su funcionamiento en una base de firmas que describen distintos perfiles de ataques. Como consecuencia solamente se detectarán amenazas descritas en la base de firmas, dejando a un lado ataques nuevos o variaciones ligeras de ataques conocidos. Dentro del área de investigación en IDS se encuentra otra estrategia de análisis que utiliza redes neuronales como parte de su funcionamiento, aunque existen distintos modelos de redes neuronales y algoritmos de entrenamiento para estas, el principal objetivo de esta investigación es plantear una nueva arquitectura junto con su algoritmo de entrenamiento para su aplicación en un IDS. Para llevar a cabo este propósito a continuación se definen los objetivos específicos de esta investigación:

- El diseño de una arquitectura recurrente con wavelets que contemple esquemas de interconexión, número de capas, número de nodos y propiedades dinámicas de la misma.
- La deducción del algoritmo que permita el proceso de aprendizaje de la red neuronal recurrente con wavelets, durante dicho proceso el algoritmo establecerá la manera en la que los parámetros libres de la red llevan a cabo su actualización en cada época.
- La construcción de rutinas que lleven a cabo el pre procesamiento de los datos de entrenamiento y prueba antes de ser presentados a la red neuronal, dado que es necesario definir cada patrón de entrada en un rango de valores específicos.

[3] El modelo *Open Systems Interconnection model (OSI)* establece un arquitectura estándar compuesta de distintas capas para dividir sistemas de comunicaciones. Dentro de este modelo, cada capa provee servicios a la capa inmediata superior.

- La aplicación del modelo desarrollado de redes neuronales artificiales para la detección, identificación y clasificación de tráfico nocivo sobre el protocolo HTTP[4] de capa de aplicación del modelo OSI.
- La simulación del modelo diseñado, llevando a cabo un análisis de los procesos de entrenamiento y prueba que permitan la obtención de indicadores estadísticos para fines comparativos con otras arquitecturas recurrentes presentadas.

1.5. Alcances del proyecto y limitaciones

- El sistema de detección de intrusos que se implementará no recolectará información de agentes distribuidos alrededor de la red, es decir que no se llevará a cabo la fusión de datos de los distintos IDS establecidos en la red.
- Este modelo de red recurrente propuesto es una mejora al modelo de red neuronal cuyas funciones de activación son sigmoides, con lo cual se espera aumentar la velocidad de convergencia e índices de detección y clasificación.
- Dentro de la investigación en modelos de IDS que utilizan la técnica de detección de anomalías se pueden dar dos fenómenos distintos: el primero es cuando las actividades anómalas que no son intrusivas son etiquetadas como intrusivas, el segundo está relacionado con actividades intrusivas que no poseen características anómalas siendo catalogadas por el IDS como tráfico de red normal. Los fenómenos descritos anteriormente pueden dar lugar a errores en la clasificación, es decir la emisión de falsas alarmas o la no detección de ataques [BAI03].
- Se tiene que llevar a cabo un pre procesamiento al conjunto de datos antes que puedan ser presentados como instancias de entrenamiento o prueba a la red neuronal [SAN05]. Dicho pre procesamiento tiene como objetivo preparar el flujo de entrada a estableciendo un tamaño fijo a cada patrón de entrada.

1.6. Organización de la Tesis

El trabajo reportado en esta tesis se organiza en 7 capítulos y 2 apéndices, a continuación se presenta una breve descripción de cada sección.

En el capítulo 2 se lleva a cabo una revisión de la literatura sobre el tema de investigación. Se da una introducción a los sistemas de detección de intrusos, su clasificación y tipos de tecnologías que utilizan. Se introducen conceptos relacionados con las redes neuronales recurrentes, propiedades, arquitecturas disponibles, algoritmos de

[4] Hypertext Transfer Protocol, es el protocolo usado en transacciones web. Este protocolo trabaja en la capa de aplicación de modelo OSI y está orientado a realizar transacciones a través del esquema petición-respuesta entre un cliente y un servidor.

entrenamiento y se describe la manera en la que se combina esta área con los sistemas de detección de intrusos.

El capítulo 3 presenta una breve introducción a la transformada wavelet, su definición, descripción matemática, elementos que la componen, descomposición y reconstrucción y procesos que lleva a cabo. Una vez cubiertos los conceptos básicos del procesamiento wavelet se presentan las redes neuronales wavelet, se da una breve introducción sobre el origen de estas, las arquitecturas básicas, características de las unidades de procesamiento y algoritmos de entrenamiento.

El capitulo 4 describe el modelo propuesto (SRWNN - MRW), se presentan distintos gráficos que proveen una descripción funcional y estructural sobre el modelo: diagrama de flujo de señal, diagrama de bloque, gráficas de flujo de señal, gráfica arquitectural. Estos elementos describen a detalle las características de la arquitectura propuesta, en la parte final de esta sección se define el algoritmo de entrenamiento acompañado de un pseudocódigo.

El capítulo 5 define del mecanismo de pre procesamiento de datos. A continuación se detallan características sobre el diseño de cada topología utilizada en la implementación, como lo son: número de capas de componen el modelo, número de unidades que componen cada capa, características de las unidades de procesamiento y descripción vectorial de sus estructuras.

En el capítulo 6 se presentan los resultados comparativos entre las distintas arquitecturas implementadas, se muestran distintas tablas y gráficos que permiten analizar los porcentajes obtenidos en distintos indicadores estadísticos para después emitir conclusiones sobre el comportamiento de cada modelo aplicado a los IDS.

En el capítulo 7 se sintetizan los resultados obtenidos y se generan las conclusiones del modelo propuesto, también se presentan posibles mejoras o áreas de oportunidad para trabajo a futuro.

En el apéndice A se dan detalles sobre la transformada wavelet que servirán como referencia al lector para profundizar sobre algunos conceptos definidos en los capítulos posteriores.

Dentro del apéndice B se ilustra el algoritmo de entrenamiento para el modelo auto recurrente con wavelets (SRWNN), se lleva cabo una presentación detallada de la arquitectura la cual incluye el desarrollo de cada uno de los pasos para la obtención del algoritmo de gradiente descendente que permita llevar a cabo el proceso de entrenamiento.