ISSN: 1692-7257 - Volumen 2 – Número 28 - 2016

Revista Colombiana de Tecnologías de Avanzada

Recibido: 23 de febrero de 2016
Aceptado: 10 de abril de 2016

DISEÑO DE UN SISTEMA MULTIAGENTES HÍBRIDO BASADO EN APRENDIZAJE PROFUNDO PARA LA DETECCIÓN Y CONTENCIÓN DE CIBERATAQUES

DESIGN OF A MULTIAGENT HYBRID SYSTEM BASED ON DEEP LEARNING FOR THE DETECTION AND DEFENSE OF CYBER ATTACKS.

PhD(c). Enrique Javier Santiago*, PhD. Jesús Sánchez Allende**

***Universidad Francisco de Paula Santander**, Facultad de Ingeniería
Grupo de Investigación GITYD, Ocaña, Norte de Santander, Colombia.
E-mail: ejsantiagoc@ufpso.edu.co
** **Universidad Alfonzo X el Sabio**, Escuela Politécnica Superior.
Ave. de la Universidad No. 1, Villanueva de la Cañada, 28691, Madrid. España
E-mail: jallende@uax.edu.

Resumen: Este trabajo es el resultado de la investigación de los autores sobre el estado del arte de la ciberseguridad a nivel mundial, y a partir de las necesidades identificadas plantea el diseño de una solución de seguridad informática resultado de la integración de dos tecnologías: los sistemas multiagentes reactivos y BDI con las redes neuronales de aprendizaje profundo para la detección y contención de ataques informáticos tradicionales y avanzados.

Palabras clave: SMA, Agentes BDI, DNN, redes neuronales, dataset.

Abstract: This work is the result of the authors' research on the state-of-the-art of cybersecurity worldwide, and based on the needs identified, the design of a computer security solution resulted from the integration of two technologies: reactive multiagent systems and BDI agents with neural networks of deep learning for the detection and containment of traditional and advanced computer attacks.

Keywords: SMA, BDI agent, DNN, neuronal network, dataset.

1. INTRODUCCIÓN

El propósito de este trabajo es proponer el diseño de una solución de ciberseguridad basada agentes inteligentes colaborativos que permitan la detección y contención de ataques avanzados contra sistemas computacionales integrados a redes TCP/IP como resultado de nuestra investigación doctoral el área de ciberseguridad. El resultado de esta investigación podría contribuir a la transformación de los tradicionales sistemas de ciberdefensa aislados en un conjunto de procesos sinérgicos integrados que actúen como un organismo digital quien a través de la coordinación de sus componentes pueda detectar y neutralizar amenazas sin asistencia humana, basando su éxito en las lecciones aprendidas.

2. ESTADO DEL ARTE DE LA CIBERSEGURIDAD

Según el informe de ciberseguridad publicado por Cisco (Cisco Systems, 2017), En el 2016 se notó una expansión en la superficie de ataque[1] gracias a

[1] Conjunto de elementos tecnológicos que pueden servir para comprometer los activos de información.

la integración de los equipos móviles, los servicios de la Nube y elementos del Internet de las cosas IoT con las redes corporativas, dándole más espacio de operación a los hackers maliciosos. Esta compañía estima que para el año 2020 el 66% de los ordenadores serán móviles y dispositivos inalámbricos que podrán ser aprovechados para llevar a cabo ataques al IoT, como ya se evidencio con la Botnet Mirai[2] el año pasado.

Por otra parte *McAfee labs* en su reporte sobre Amenazas (*McAfee Labs*, 2016) informa que en el 2016 hubo un incremento en los ataques de *ransomware*[3], en el número de variantes de sus cepas, y en la fuga de información en las empresas a pesar de una adopción creciente de sistemas de prevención de perdida de datos –DLP por la inefectividad de las técnicas de identificación de activos basadas en expresiones regulares, también dice su informe que entre el 20% y el 40% de estos robos son perpetrados por personal interno como empleados, contratistas y socios de negocios.

2.1 Fallas de los controles actuales

Algunas de las limitaciones de las soluciones de ciberseguridad provienen de tener un único punto de fallo, recursos de procesamiento limitados y la imposibilidad de adaptarse a los cambios de estrategia evidenciados en los ataques de última generación.

El vicepresidente de la empresa de ciberseguridad *Symantec* Brian Dye expreso en entrevista al *Wall Street Journal* (Danny Yadron, 2014), que los antivirus ya no son efectivos debido a la evolución creciente de las técnicas de *hacking* y demás ataques informáticos. En ese momento informo que el porcentaje de detección efectiva estaba en un 45% y que por esto su compañía cambio el foco de las soluciones de seguridad ofrecidas.

Cisco afirma que debido a la complejidad de los ataques informáticos actuales, las empresas están usando cada vez más tecnologías de varios fabricantes al tiempo, haciendo más compleja la gestión de la ciberdefensa, y sugiere que las empresas hagan uso de soluciones que sean integrales, que simplifiquen la seguridad de sus operaciones y que confíen más en la automatización. Dicen también que este enfoque ayudará a reducir los gastos operacionales, aliviar la carga sobre el personal de seguridad y ofrecer mejores resultados de seguridad.

En su informe *McAffe* mas reciente hace referencia al aprendizaje automático y a su aplicación práctica en ciberseguridad para mejorar la detección de amenazas.

3. SOLUCIÓN PROPUESTA

3.1 Descripción general

La solución propuesta forma parte de la investigación (Santiago y Sánchez, 2017) realizada en el área de la seguridad informática con respecto a las principales amenazas que actualmente afectan a los activos de información de las empresas y a las oportunidades de mejora que presentan las diferentes tecnologías de detección y contención de ciberataques que son usadas actualmente por las organizaciones.

Se tomaron como referencia algunas investigaciones previas en el área de sistemas multi-agentes (Bonfante y Castillo, 2014), de IDS basado en agentes (Parra y Herrera, 2013) y el uso de redes neuronales artificiales (Gualdròn y Durán, 2014) entre otras.

En este trabajo se propone abordar los fallos en la detección y tratamiento de ataques de última generación a través del diseño de un Sistema Multi Agentes hibrido con capacidad de aprendizaje Automático profundo[4].

3.2. Arquitectura de la solución

Considerando las limitaciones de los sistemas de seguridad actuales, esta propuesta consiste en un sistema de ciberseguridad distribuida multicapas y escalable compuesto por un conjunto de agentes con capacidad adaptativa que se apoyan en un sistema de aprendizaje automático avanzado.

Los agentes que conforman la solución, están divididos en 3 capas:

- Capa de monitorización.
- Capa de análisis.
- Capa de supervisión.

[2] Red de bots del "IoT" que afecto el servicio de DynDNS en el 2016.
[3] Secuestro de activos información con el uso de cifrado.
[4] Evolución del machine learning

En la figura siguiente se ilustra la distribución de los agentes en las capas antes mencionadas.

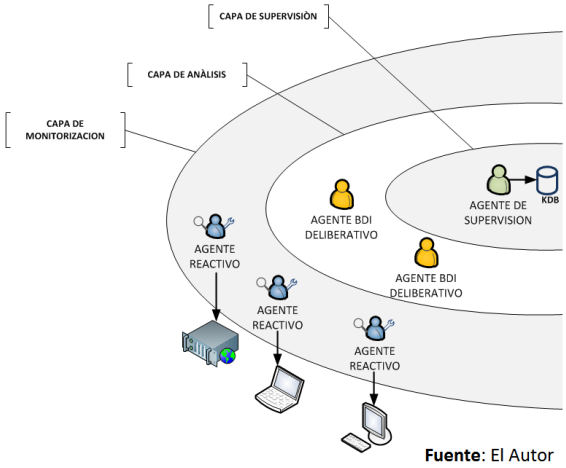

Fig. 1. arquitectura de la solución SMA[5].

La primera capa tiene como objetivo obtener los parámetros necesarios para facilitar la identificación de acciones maliciosas que podrían considerarse ataques contra los activos de información de cualquier empresa.

La segunda capa lleva a cabo las actividades de procesamiento de los parámetros obtenidos por los agentes de monitorización que son necesarios para determinar si los ordenadores en evaluación están enfrentado una situación de riesgo.

La tercera capa es la encargada de realizar el registro de actividades de los agentes de las capas anteriores y su posterior almacenamiento persistente en una base de datos de la cual pueden extraerse reportes referentes a los eventos e incidentes de seguridad detectados.

3.2.1 Capa de monitorización

Está compuesta por un grupo de agentes reactivos distribuidos sobre el conjunto de ordenadores que deben protegerse. Estos agentes están compuestos por un módulo de lectura del tráfico de las interfaces de red activas de cada sistema monitorizado, un módulo pre-procesador de los diferentes flujos[6] de tráfico leídos en tiempo real, un módulo analizador de tráfico, un módulo forense, un módulo de entrada/salida y defensa más un módulo de comunicaciones a través del cual el agente interactúa con los otros agentes de la solución.

Fig. 2. Arquitectura de un agente reactivo.

En la figura anterior se muestra la arquitectura del Agente reactivo propuesto; el módulo de lectura de tráfico se inicia como parte del "comportamiento" de este agente, e identifica las interfaces de red del ordenador, de las cuales selecciona la que será monitorizada.

3.2.1.1 Descripción de componentes de los agentes reactivos

3.2.1.1.1 Módulo de lectura de tráfico

Este componente inicia con la detección de las interfaces de red del ordenador. De cada una extrae principalmente la marca, serie, identificador asignado por el sistema operativo, dirección de hardware, dirección IP y mascara de red. Luego crea una colección con las interfaces detectadas e inspecciona cuales están activas y en uso para finalmente seleccionar la que será monitorizada.

Posteriormente se cargan en memoria las librerías de código necesarias para iniciar el proceso de lectura de tráfico, que consiste en la interceptación del flujo de bytes que es decodificado por la interfase de red física del ordenador y es enviado posteriormente a un área de la memoria RAM controlada por el agente, donde la copia del tráfico será almacenada.

Las rutinas de código de este módulo crean instancias de cada trama, paquete, datagrama y segmento que entra o sale de la interfase de red, que finalmente tendrá como atributos los parámetros de la cabecera y del área de datos de cada estructura de mensaje, facilitando la extracción de valores de elementos necesarios para su posterior inspección y análisis.

[5] Sistema multi-agentes

[6] Conjunto de paquetes que tienen las mismas características

*3.2.1.1.2 Módulo de pre-procesamiento de tráfico***.**

Una vez el módulo de lectura tiene el tráfico encapsulado en objetos, el agente hace uso del "módulo de pre-procesamiento de tráfico" a través del cual extrae la información de la cabecera y del área de datos del flujo leído. Con el fin de facilitar la detección de las amenazas, se hace necesario identificar las características particulares de los ataques que se pretende detectar. Estas características pueden ser: valores de los campos en las cabeceras de los paquetes, datos específicos encapsulados en su área de carga útil, combinaciones de ambos e incluso tasa de bits de paquetes del mismo tipo que puedan servir para identificar un patrón. Con toda esta información se construye un arreglo (DATASET) que será usado posteriormente por el analizador de tráfico básico, será enviada a través de la red a la segunda capa de la solución compuesta por un conjunto de agentes inteligentes deliberativos BDI encargados de procesar la información recibida.

La propuesta considera la detección de las siguientes acciones:

- Escaneo de puertos por "XMAS SCAN".
- Escaneo de puertos por "NULL SCAN".
- Ataque de ARP *Spoofing*.

A continuación se explican las consideraciones y el patrón buscado por el analizador para identificar cada una de las acciones ilegales previamente relacionadas.

Escaneo de puertos con la técnica XMAS-SCAN

Esta actividad en sí misma no representa un ataque, pero su ejecución es considerada ilegal porque generalmente es ejecutada dentro de muchos ciberataques con el fin de identificar la presencia de los servicios y números de puerto en los ordenadores víctima.

El proceso consiste en el envío a través de la red de varios segmentos TCP con las banderas FIN, PSH, URG activas al tiempo, en el campo denominado "*bit code*" dentro de la cabecera TCP a un rango de puertos lógicos del ordenador víctima. De estar abierto alguno de los puertos interrogados, el servicio que escuche a través de este, intentará identificar la solicitud del *host* remoto a partir de la identificación de las banderas activas. A continuación se muestra la composición de un segmento TCP construido para realizar un XMAS *scanning* en el que puede apreciarse las banderas activas.

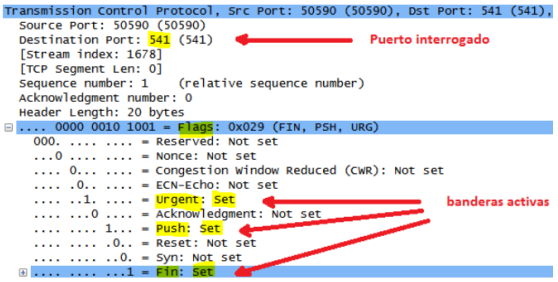

Fig. 3. TCP de un XMAS Scanning

Una vez que servicio destino recibe el trafico XMAS en el ordenador víctima, determinara que es una solicitud invalida, pero ninguna implementación actual del protocolo TCP contempla esta combinación de banderas. (Verificado en el RFC 793 para el protocolo TCP).
La respuesta del ordenador victima depende del sistema operativo que tenga instalado. Unix/Linux/MacOS en cualquiera de sus versiones responde con un segmento TCP con la bandera RST activa, mientras que los sistemas operativos Microsoft Windows no envían respuesta alguna.
Estas reacciones le permiten al agresor identificar el estado del puerto evaluado y por ende el servicio asociado a este.

Una vez extraídos los datos, el "preprocesador de tráfico" construye el siguiente DATASET[7] que posteriormente es enviado al analizador de tráfico simple.

Ataque de ARP Spoofing:

Otro ataque que podría detectar la primera capa el sistema multi-agentes es el envenenamiento de las tablas ARP de los ordenadores en redes Ethernet conmutadas, muy utilizado para materializar la intercepción de tráfico como parte de una técnica conocida como MiTM[8].

Este ataque va dirigido al protocolo de resolución de direcciones ARP, que sirve de interfase entre el protocolo de nivel de enlace "*Ethernet*" (o alguna de sus variantes) y el protocolo de nivel de red "IP" y que está presente en la mayoría de los ordenadores de las empresas de la actualidad; la agresión pretende crear una entrada falsa en la tabla ARP a través de continuos mensajes ARP gratuitos en la que se asocia la dirección de hardware MAC de la interfase de red del agresor

[7] Colección de datos que incluye la representación estructurada del tráfico a procesar

[8] Ataque denominado de hombre en el medio

con la dirección IP del *host* legítimo que quiere suplantarse tal como se muestra a continuación:

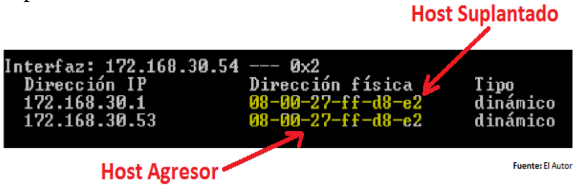

Fig. 4. Tabla ARP del host víctima.

El fin de la acción maliciosa es poder interceptar el tráfico que se intercambia entre dos procesos típicamente cliente y servidor que se encuentran distribuidos en dos ordenadores diferentes.

Con esta acción por ejemplo, todo el tráfico que es enviado desde el cliente al servidor será entregado al ordenador del agresor incluso si va cifrado, si este hace uso de herramientas como *Bettercap*[9], sslstrip2.0. Incluso podría interceptarse el tráfico asociado al intercambio de llaves criptográficas y tener acceso al contenido en "claro" de los mensajes intercambiados.

A continuación puede apreciarse los campos de la cabecera Ethernet y ARP que son relevantes para la identificación de un ataque de envenenamiento de ARP.

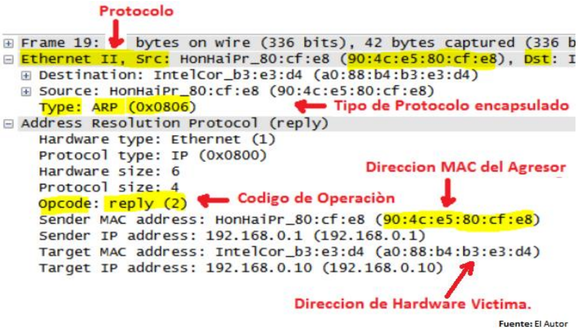

Fig. 5: Metadatos de tráfico relevantes para la detección.

Muchas herramientas de *hacking* como Cain & Abel, Ettercap y ARPSpoof mantienen el ataque activo a través del envió continuo en el tiempo de una gran cantidad de mensajes ARP gratuitos evitando así que la información falsificada de la tabla ARP de la víctima sea corregida. De manera que el tráfico de este tipo de ataque podría verse en un sniffer[10] tal como se muestra a continuación.

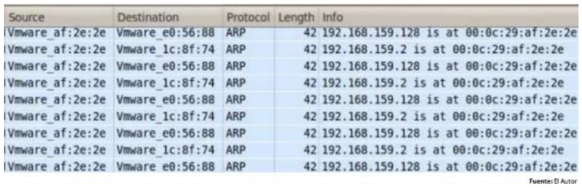

Fig. 6. Tráfico continuo de mensajes ARP reply de un ataque de ARP Spoofing.

Para que los agentes reactivos puedan detectar el tráfico malicioso, el flujo capturado por el módulo de lectura de tráfico debe ser enviado al pre-procesador que debe leer flujo por flujo el tráfico recibido e ir identificando tramas Ethernet y mensajes ARP con destino al ordenador monitorizado, extrayendo continuamente los valores de los siguientes campos significativos:

- Ethernet Type: (0x806) ARP
- ARP Opcode: (2) reply y (1) request
- ARP Sender Mac Address (Agresor)
- ARP Target MAC Address (Victima)

También debe contabilizar la cantidad de tramas con los mismos valores en los campos antes mencionados y su delta tiempo entre tramas iguales, para determinar si estas forman parte de una inundación de este tráfico, como también es requerido que valide que los mensajes ARP *reply* no son resultado de la respuesta de un ARP *request* enviado previamente por el ordenador monitorizado.

Con los datos obtenidos, el "pre-procesador de tráfico" construye el siguiente DATASET que posteriormente es enviado al analizador de tráfico simple.

3.2.1.1.3 Módulo analizador de tráfico simple

Una vez recibidos los DATASET, este módulo se encarga de procesar sus datos e identificar la presencia de un patrón de ataque conocido. Por ejemplo, para el caso del escaneo de puertos XMAS, el proceso de detección se basa en realizar una búsqueda dentro del arreglo, de los registros de trafico TCP con las banderas URG, PSH y FIN activadas para determinar con éxito que ese flujo de tráfico coincide con el ataque evaluado.

Con respecto a la detección del ataque ARP *Spoofing* mencionado con anterioridad; una vez recibido el DATASET, se buscan los registros de tramas Ethernet con mensajes ARP reply encapsulados, que sean gratuitos y continuos en el tiempo para determinar que el sistema se encuentra ante la presencia de un ataque de MiTM del tipo explicado anteriormente.

[9] Herramienta que junto con Sslstrip facilitan el acceso a tráfico cifrado.
[10] Herramienta para capturar y analizar trafico

3.2.1.1.4 Módulo forense
Una vez finaliza el proceso de análisis de tráfico y se confirma la detección de una amenaza, el agente activa las funciones forenses que llevan a cabo la recolección de los datos de la memoria volátil que incluyen principalmente a la tabla de conexiones NetBIOS sobre TCP/IP, la tabla de enrutamiento, la tabla ARP, el registro de conexiones activas, la lista de procesos en ejecución, el listado de unidades de red en uso, las sesiones de usuario activas, etc. Toda esta información es considerada vital, junto con las imágenes de los discos duros y demás unidades de almacenamiento persistente por los analistas forenses dentro de un proceso de investigación.

Ante un incidente de ciberseguridad muchos usuarios y organizaciones por temor a la agresión interrumpen la energía de los ordenadores eliminando mucha de la evidencia contenida en los datos volátiles.

Al finalizar la recolección de la información volátil, se calcula la huella dactilar con el algoritmo SHA1 con el fin de legalizar la extracción y luego junto con su HASH es enviada a través de mensajes a los agentes de la capa de supervisión para su registro y almacenamiento.

3.2.1.1.5 Módulo de defensa
El módulo de defensa es un componente del agente reactivo encargado de interactuar con los diferentes controles del *host* monitorizado que podrían ayudar a contrarrestar la acción agresiva, como el Firewall del sistema Operativo y el *Antimalware*. Su rutina de código se ejecuta al finalizar la ejecución del módulo forense como resultado de la detección positiva de un ciberataque por parte del módulo analizador de tráfico o de los agentes deliberativos BDI de la segunda capa del SMA. Su implementación ejecuta las acciones necesarias para contener la agresión dependiendo del tipo de ataque informado en el resultado del análisis.

Si el ataque detectado es un escaneo de puertos, entonces será bloqueado el tráfico generado por el agresor.

Si es un ataque de ARP *Spoofing*, procede a bloquear el acceso a cualquier servicio del *host* desde la dirección del agresor, luego limpiar la tabla ARP y crea un registro estático en esta tabla con la información correcta.

Si se detectó alguna conexión saliente fraudulenta, procede a identificar el proceso local que tiene la conexión establecida, se cierra la conexión y se invoca al antimalware para que evalúe su estado.

3.2.1.1.6 Módulo de comunicaciones
Al finalizar la ejecución de las actividades de defensa, el agente del ordenador atacado usa su módulo de comunicaciones para informar a sus vecinos sobre el incidente ocurrido, enviando un mensaje con la descripción del tipo de ataque sufrido y la identificación del agresor, esto con el fin de que cada uno ejecute las acciones de defensa necesarias para evitar una agresión futura proveniente de la misma fuente (Duran & Iturriago, 2012).

Este módulo es uno de los elementos fundamentales para el funcionamiento de la solución propuesta ya que, a través de este, los agentes pueden interactuar de forma colaborativa intercambiando mensajes de información o de control entre pares, sin importar la capa a la que pertenezcan.

3.2.1.1.7 Módulo de registro de actividades
Cada evento detectado por los sensores de los agentes al igual que la ejecución de sus tareas es registrado en una bitácora local que es enviada posteriormente a los agentes de la capa de supervisión para su almacenamiento con el fin de facilitar el cálculo de estadísticas y reportes de actividades del sistema y que de ser necesario pueda hacerse trazabilidad de las acciones de cada componente de la solución multiagentes.

3.2.2 Capa de Análisis
Los agentes inteligentes son los encargados de recibir la solicitud por parte de los agentes reactivos del procesamiento del tráfico interceptado, junto con los DATASET que por su complejidad no pueden ser procesados con las reglas del analizador de trafico simple de la capa de monitorización. Estos agentes son el núcleo de la solución multiagentes, su estructura es un poco más compleja.

Los agentes inteligentes de esta capa están compuestos por un módulo de comunicaciones a través del cual interactúan con los agentes reactivos, un módulo de aprendizaje adaptativo compuesto por dos redes neuronales multicapa y un módulo de clasificación del tráfico y un módulo de registro de actividades. La arquitectura de estos agentes inteligentes se muestra en la figura siguiente:

Fig. 7. Arquitectura de un agente inteligente BDI.

3.2.2.1 Módulo de comunicaciones
Este módulo sirve de interface entre el agente inteligente y los otros agentes de todas las capas. A través de este se reciben los DATASET, se envía el registro de actividades a la capa de supervisión y se ordena a los agentes reactivos realizar las acciones de defensa en el momento de la detección de una agresión.

3.2.2.2 Módulo de aprendizaje automático
Una vez es recibido el arreglo de patrones de comportamiento observado por la primera capa, los agentes inteligentes (Russell & Norvig, 2005). Inician el proceso de análisis a través del módulo aprendizaje automático compuesto por dos redes neuronales multicapa de aprendizaje profundo - DNN (Mansanet Sandín, J., 2016), cada perceptron cuenta con una función de activación diferente que va a permitir que la clasificación tenga un mayor grado de confiabilidad.

El módulo de aprendizaje automático tiene como objetivo clasificar el tráfico monitorizado como normal, ilegal o sospechoso de acuerdo a los datos recibidos en el DATASET.

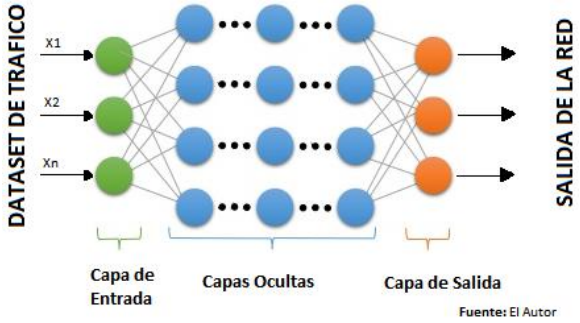

Fig. 8. Capas de las Redes neuronales usadas por los Agentes Inteligentes.

Las dos (2) DNN son discretas y están compuestas por perceptrones con (5) cinco capas de neuronas: una de entrada, tres ocultas y una de salida. La distribución de la cantidad de neuronas es el resultado del número de elementos *Xi* de los vectores de entrada y del cálculo del número de neuronas efectivas requeridas por capa considerando los umbrales de sobre entrenamiento e inefectividad del proceso de clasificación.

Cada red neuronal artificial hace uso de una función de activación no lineal diferente que le permite al sistema multiagentes realizar con mayor precisión la clasificación del tráfico. Las funciones de activación elegidas fueron la función hiperbólica tangencial y la sigmoide.

3.2.2.3 Módulo de clasificación de tráfico
Al finalizar el cálculo de ambas redes neuronales, los valores resultantes son pasados como parámetros al módulo de clasificación de tráfico el cual comparará ambos resultados y decidirá si el tráfico analizado es normal, representa un ataque o no se puede determinar. En caso de que el resultado de las RNA no sea coincidente, el clasificador asumirá que el tráfico es sospechoso.

3.2.2.4 Módulo de registro de actividades
Tal como ocurre con los agentes de la primera capa, cada tarea realizada por cada módulo de los agentes BDI es registrada en la bitácora local y enviada posteriormente a los agentes de la capa de supervisión para su registro y almacenamiento persistente.

3.2.3 Capa de Supervisión
La última capa de la solución o capa de supervisión, está compuesta por la instancia de uno o más agentes reactivos que tienen como función recibir flujos de información referentes al registro de actividades de cada agente activo de todas las capas de la solución.

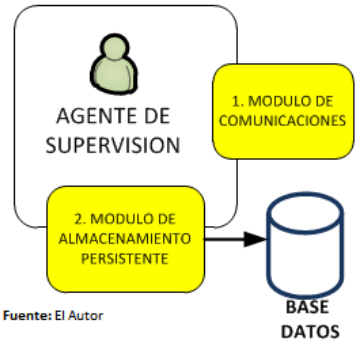

Fig. 9. Arquitectura de un agente de supervisión.

Estos agentes cuentan con un módulo de comunicaciones a través del que reciben los datos que deben registrar y un módulo de almacenamiento persistente que realiza el guardado de la información en una base de datos centralizada que facilita la generación de reportes.

3.2.3.1 Módulo de comunicaciones
Al igual que el resto de los agentes, estos cuentan con un módulo de comunicaciones que les sirve de interface para interactuar con sus pares en las otras capas principalmente para la recolección de la información de actividades del sistema multiagentes.

3.2.3.2 Módulo de almacenamiento persistente
Este módulo se encarga del almacenamiento de la información recibida, en una base de datos centralizada que posteriormente puede ser usada como fuente de información para construir reportes referentes a todos los eventos de ciberseguridad detectados por el sistema multiagentes, hacer cálculos estadísticos e incluso predicciones, haciendo uso de técnicas de minería de datos, aunque esta última fase no está considerada en el proyecto pero podría formar parte de trabajos futuros.

4. VALIDACIÓN DEL MODELO EN UN ENTORNO DE PRUEBAS

4.1 Descripción especifica del piloto de pruebas

El diseño de la solución de ciberseguridad propuesta será validado en un entorno de laboratorio TESTBED a través de la implementación de un "piloto" del mismo con el lenguaje de programación Java, usando la Plataforma de construcción de agentes JADE, el Framework de implementación de agentes deliberativos DBI llamado BDI4Jade, el paquete JPCAP para la gestión de las interfaces de red y captura de trafico TCP/IP, y un conjunto de librerías de redes neuronales de aprendizaje profundo implementadas por los autores.

Las redes neuronales usadas en la implementación del piloto, por ser de aprendizaje supervisado requieren de un proceso previo de entrenamiento a través de un "*dataset*" que será construido a partir del registro de tráfico obtenido de la ejecución de un conjunto de ataques informáticos de última generación disparados contra unos ordenadores de pruebas.

Para la realización de las pruebas empíricas, y con el fin de acotar las mismas, el autor seleccionó el ataque denominado "spear phishing" (Trend Micro,2012) o "phishing dirigido" considerado como un ataque de ingeniería social basado en computación; como uno de los ataques usados para probar la efectividad de la segunda capa de agentes inteligentes de la solución propuesta, ya que este es usado con mucha frecuencia por hackers maliciosos contra los empleados de las empresas después de haber obtenido información personal en la fase de reconocimiento durante la ejecución de ataques persistentes avanzados, que no solo afectan las finanzas de muchas organizaciones sino también la de muchas personas que hacen transacciones a través de la internet.

4.2 Ventajas de la solución propuesta

El sistema multiagentes propuesto ofrece a las organizaciones las siguientes ventajas sobre las soluciones existentes:

- Aprendizaje Adaptativo que le permite evolucionar ante los cambios de estrategia de los ataques informáticos.
- Mayor confiabilidad en la detección.
- Capacidad distribuida.
- Recuperación rápida de errores
- Tolerancia a fallos.
- Cubre nuevos tipos de ataque.

5. CONCLUSIONES

La adopción de tecnología en los procesos de negocio de las empresas de todo tipo está en crecimiento continuo gracias a los beneficios que trae consigo, pero también gracias a esto, las compañías que no gestionan adecuadamente la seguridad de sus activos de información, más temprano que tarde se ven involucradas en incidentes de seguridad que terminan afectan su operación.

Muy diferente a lo que se esperaría en una época de crecimiento y uso común de sistemas DLP, *antispam*, *firewalls* de siguiente generación y sistemas de prevención de intrusiones entre otros; es el incremento bastante significativo en la cantidad de ciberataques exitosos a organizaciones de todos los sectores de la economía. Todo esto como resultado de la expansión de la superficie de ataque que obedece a la integración de nuevas tecnologías como el Internet de las Cosas, la Nube, los móviles y los *weareables* a las redes corporativas y al cambio de estrategia de los

agresores que ahora trabajan de forma organizada con cada vez más modernas y mejores herramientas que impactan incluso en los mismos sistemas de control gracias a sus vulnerabilidades y a la falta de adaptabilidad que tienen ante las nuevas estrategias de agresión.

Los ciberataques, el malware y las técnicas de *hacking* han evolucionado extendiéndose prácticamente a cualquier tipo de tecnología conectada a la *internet*, pero muchos sistemas de ciberdefensa utilizan técnicas estáticas ante los cada vez más cambiantes ataques informáticos. Hoy en día el uso constante de técnicas de ingeniería social se aprovecha de vulnerabilidades en las personas que se vuelven determinantes a la hora de materializar una agresión contra los activos de información digital y que son difíciles de detectar por la tecnología de defensa tradicional. Por esta razón como resultado de un proceso de investigación a conciencia, en este artículo se propone una solución de ciberseguridad escalable, distribuida, tolerante a fallos y con la capacidad de adaptarse a los cambios de técnicas y estrategias usadas por los ciberagresores. El sistema de ciberseguridad propuesto se basado en agentes inteligentes que se distribuyen entre los ordenadores de las empresas podrían mejorar la tasa de éxito de la detección de ataques tanto tradicionales como de última generación reduciendo la participación humana tanto en la detección como en la contención de amenazas digitales.

REFERENCIAS

Santiago E., Sánchez J. (2017). Riesgos de ciberseguridad en las empresas, Universidad Alfonzo X el Sabio, Madrid, España.

Trend Micro, (2012): Spear Phishing, TrendLabs, Cupertino, California.

CISCO Systems, (2017): Informe de Ciberseguridad de Cisco 2017, Santa Cruz, Estados Unidos de América.

Mcfee Labs, (2016): Informe sobre amenazas de las empresas, Estados unidos de América.

Danny Yadron, Wall Street journal (2014): Efectividad de los antivirus, según vicepresidente de Symantec *https://www.wsj.com/news/articles/SB10001424052702303417104579542140235850578* (Consultado el 19 de Octubre del 2016)

Russell, S. Norvig, P. (2005). Inteligencia artificial, un enfoque moderno, Ed. Pearson, segunda edición, Madrid.

Durán Acevedo Christian M, Iturriago Ali Xavier. (2012). Automatización de un Sistema de Suministro de Agua Potable a Través de la Tecnología Zigbee. Revista Colombiana de Tecnologías de Avanzada, Vol. 1, No. 19, pp. 36-42.

Tomás, V. and García, L. (2005). "A Cooperative Multiagent System For Traffic Management And Control", AAMAS Utrecht, Netherlands.

Mansanet Sandín, J. (2016). Contributions to Deep Learning Models [Tesis doctoral]. Universitat Politècnica de València. doi:10.4995/Thesis/10251/61296.

Bonfante, Castillo (2014). "Integración de sistema multi-agente, ontologías y procesos de negocios como marco tecnológico de la estrategia "gobierno en línea". Revista Colombiana de Tecnologías Avanzadas, ISSN: 1692-7257, Vol. 1, No. 23, 2014.

Parra, C., Herrera, (2013). "Aplicación de los sistemas de detección de intrusos y la tecnología de agentes en el monitoreo inteligente de redes de datos", Revista Colombiana de Tecnologías Avanzadas, ISSN: 1692-7257, Vol. 2, No. 22, 2013.

Gualdrón, O., Durán, C. (2014). "Implementación de un modelo neuronal en un dispositivo hardware (FPGA) para la clasificación de compuestos químicos en un sistema multisensorial (nariz electrónica)", Revista Colombiana de Tecnologías Avanzadas, ISSN: 1692-7257, Vol. 2, No. 24, 2014.