UNIVERSIDAD DE BURGOS
ESCUELA DE DOCTORADO

TESIS DOCTORALES

TÍTULO: "TÉCNICAS DE CLUSTERING COMO COMPLEMENTO A MOVICAB-IDS PARA LA DETECCIÓN DE INTRUSIONES"

AUTOR: SÁNCHEZ ARÉVALO, RAÚL

PROGRAMA DE DOCTORADO: TECNOLOGÍAS INDUSTRIALES E INGENIERÍA CIVIL

FECHA LECTURA: 23/03/2018

HORA: 11:30

CENTRO LECTURA: ESCUELA POLITÉCNICA SUPERIOR. SALÓN DE GRADOS. CAMPUS LA MILANERA

DIRECTOR: ÁLVARO HERRERO COSÍO

TRIBUNAL: EMILIO SANTIAGO CORCHADO RODRÍGUEZ
ÁNGEL ARROYO PUENTE
JOSÉ LUIS CALVO ROLLE
LETICIA ELENA CURIEL HERRERA
HÉCTOR QUINTIÁN PARDO

RESUMEN:

La naturaleza siempre cambiante de las tecnologías y estrategias con las que se desarrollan es uno de los aspectos más perjudiciales de los ataques e intrusiones que sufren los sistemas de información y sus redes. Esta circunstancia aumenta la dificultad de conseguir una protección exitosa. Por esa razón, entre otras, los Sistemas de Detección de Intrusiones (IDS por sus siglas en inglés) se han convertido en un activo esencial de la infraestructura de seguridad informática en la mayoría de las organizaciones.

En el contexto de las redes de computadores, un IDS se puede definir en términos generales como una herramienta diseñada para detectar patrones sospechosos que pueden estar relacionados con un ataque a un sistema. La Detección de Intrusiones (ID) es por lo tanto, un campo que se centra en la identificación de los intentos de ataque en curso en un sistema informático (Host IDS - HIDS) o de red (Network IDS - NIDS).

Dentro del campo de los IDS se propuso MOVICAB-IDS (MObile VIsualisation Connectionist Agent-Based IDS) como un sistema de Inteligencia Artificial Híbrido (IAH) basado en un sistema multiagente que incluye agentes deliberativos capaces de aprender y evolucionar con su entorno. Estos agentes combinan un modelo neuronal proyeccionista basado en aprendizaje no supervisado y el paradigma de Razonamiento Basado en Casos. MOVICAB-IDS supervisa la actividad de la red para identificar eventos intrusivos mediante la combinación de diferentes paradigmas de la Inteligencia Artificial (AI). Con ellos se lleva a cabo una visualización del tráfico de red con el objetivo de llevar a cabo la detección de intrusiones, analizando los datos que viajan por ella. Su objetivo principal es proporcionar al personal de seguridad una visualización sintética e intuitiva del tráfico de la red para facilitar la detección de intrusiones y la supervisión de la

actividad en ésta.

Uno de los principales inconvenientes de MOVICAB-IDS es su dependencia del personal de seguridad ya que no puede proporcionar una alarma automática. Por ello este sistema exige que el administrador de red analice las proyecciones proporcionadas y decida si entre los datos presentados, alguno muestra un comportamiento anómalo.

La presente tesis pretende proporcionar una extensión a MOVICAB-IDS a través del uso de técnicas de agrupamiento (clustering en inglés) de forma que se resuelva la dependencia con el procesamiento humano puesto que éste puede fallar, incluso ante situaciones claras de ataque, debido a las limitaciones del procesamiento humano cuando se analizan visualmente grandes cantidades de datos.

El agrupamiento es la clasificación no supervisada de patrones (observaciones, elementos de datos, o vectores de características) en grupos (clusters) basándose en su similitud. El problema del agrupamiento se ha tratado en muchos contextos y por investigadores de muchas disciplinas, lo cual refleja su gran atractivo y utilidad como uno de los pasos en el análisis exploratorio de datos.

En la presente tesis se han aplicado exitosamente técnicas de agrupamiento a los datos empleados por MOVICAB-IDS para ver cómo éstas pueden contribuir a la detección de intrusiones e incluso ayudar a que MOVICAB-IDS sea capaz de proporcionar una respuesta automática.

A lo largo del trabajo de investigación realizado, se han validado diferentes técnicas de agrupamiento tanto sobre los datos recogidos directamente de la red (información extraída de las cabeceras de los paquetes), como sobre las proyecciones de esos mismos datos que genera MOVICAB-IDS tras su análisis. De esta manera se puede observar y contrastar la validez de la solución propuesta como extensión de MOVICAB-IDS, así como comprobar la aplicación de las técnicas de agrupamiento en el análisis de intrusiones.

Las técnicas de agrupamiento utilizadas son representativas de los dos principales tipos existentes (particionales y jerárquicas) para comprobar su rendimiento con los datos indicados. Se han aplicado k-means en el caso de los métodos particionales y el método aglomerativo para el caso de los métodos jerárquicos. Por otra parte, las técnicas de agrupamiento llevan a cabo el agrupamiento de los datos basándose en su proximidad, que se mide usualmente mediante una función de distancia definida entre cada par de datos. Existe una gran variedad de medidas de distancia, algunas de las cuales han sido utilizadas también, para las distintas técnicas de agrupamiento.

Con lo anteriormente indicado, la presente tesis analiza la posibilidad de extender MOVICAB-IDS mediante técnicas de agrupamiento de forma que suponga un avance en la dotación de su respuesta automática, así como aumentar su precisión y rapidez. Con todo ello se pretende afrontar la detección precoz de situaciones anómalas desconocidas hasta el momento (“0-day attacks”) mientras éstas están sucediendo.