Universitat Oberta de Catalunya (UOC)

Máster Universitario en Ciencia de Datos (*Data Science*)

TRABAJO FINAL DE MÁSTER

Área: Minería de datos y Machine Learning

Detección de anomalías en entorno del Internet de las cosas

Autor: Gonzalo Pedro Mellizo-Soto Díaz

Tutor: Carlos Hernández Gañán

Profesor: Jordi Casas Roma

Madrid, 23 de marzo de 2019

Copyright

FICHA DEL TRABAJO FINAL

Título del trabajo:	Detección de anomalías en el entorno del Internet de las cosas
Nombre del autor:	Gonzalo Pedro Mellizo-Soto Díaz
Nombre del colaborador/a docente:	Carlos Hernández Gañán
Nombre del PRA:	Jordi Casas Roma
Fecha de entrega (mm/aaaa):	06/2019
Titulación o programa:	Máster Universitario en Ciencia de Datos
Área del Trabajo Final:	Minería de datos y Machine Learning
Idioma del trabajo:	Español
Palabras clave	Machine Learning, IOT, Anomaly Detection

Cita

"Nuestro lema es: más humanos que los humanos"

Eldon Tyrell, *Blade Runner*

Agradecimientos

TO BE DEFINED

Si se considera oportuno, mencionar a las personas, empresas o instituciones que hayan contribuido en la realización de este proyecto.

Abstract

In recent years the amount of connected devices has greatly increased, with an increasing number of applications in the industry each day. This devices can be subject of attacks causing instability or data leaks that can be dangerous both for the users and the enterprises, in order to avoid or confront them, security and early detection are becoming a must in a connected world. The focus is the monitoring and detection of the attacks in Internet of Things devices using state of the art Machine Learning techniques. Models such as SVM, DBScan or Isolation Forests have been used and assembled in order to identify with a better accuracy when an attack is happening. With this assembly, attack detection has increased up to 15 % comparing to traditional methods and individual model usages and times have been considerably reduced. An active use of Machine Learning models has shown a great improvement at anomaly detection by securing the devices and decreasing the reaction times when facing attacks.

Durante los últimos años se encuentra una creciente cantidad de dispositivos conectados entre sí, cada vez con más aplicaciones en la industria. Estos dispositivos pueden ser atacados y provocar inestabilidad o una fuga de datos, por lo tanto la protección y la pronta detección de ataques y/o anomalías es vital en un mundo cada vez más conectado. El objetivo es la monitorización y detección de estos ataques en dispositivos del *Internet of Things* utilizando técnicas del estado del arte de Machine Learning para su detección y poder así responder con una mayor rapidez a los ataques. Para la detección se han utilizado modelos estadísticos, como SVM, DBScan o Isolation Forests, que en su conjunto permitan identificar con mayor precisión cuando se está produciendo un ataque. El conjunto de la clusterización con la clasificación de puntos anómalos muestra una mayor robustez, frente al uso individual de cada uno de los modelos aumentando la detección en hasta un 15 %. Se demuestra cómo el uso de los modelos permite proteger los dispositivos y mejorar la seguridad al disminuir los tiempos de reacción frente a los ataques.

Palabras clave: Machine Learning, IOT, Anomaly Detection

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

1.1. Contexto y justificación del Trabajo

Cada vez se encuentran más dispositivos conectados entre sí, no solo en la industria, sino también en los hogares, esta conexión entre dispositivos los hace vulnerables a ataques informáticos que pueden afectar en gran medida a los usuarios, no solo pueden provocar un mal funcionamiento de los mismos, sino que también puede provocar la fuga de datos de distinta sensibilidad. La previsión en los futuros años es estar cada vez más conectados y una pronta detección de los ataques puede ayudar a evitar los problemas derivados de los mismos, mediante una pronta reacción o haciendo frente a la previsión de un ataque.

Actualmente se utilizan distintas medidas de seguridad como control de acceso físico al dispositivo, encriptación de datos, firewalls, securización de red, etc... Sin embargo, en muchos casos la detección de anomalías aplica un umbral estacionario, provocando que en el caso de un ataque ya sea demasiado tarde para reaccionar o no se sea capaz de identificar patrones extraños previos al ataque. Con la aplicación de nuevas técnicas se pretende mejorar los tiempos de reacción e incluso predecir cuándo puede suceder un ataque.

1.2. Explicación de la motivación personal

La razón principal de la elección del proyecto es la posibilidad de aprender y utilizar técnicas de Machine Learning en un sector desconocido que permita diversificar conocimientos. Todo se encuentra cada vez más conectado e investigar cómo clasificar cuando se está produciendo un ataque permite profundizar en conocimientos de seguridad y aplicarlos en un problema real.

La aplicación de técnicas en un problema real ayuda a comprender mejor el uso de las herramientas y el porqué y cuándo se deben de utilizar. De este modo, se añade un nuevo conocimiento que aportar al ámbito profesional y puede utilizarse también en un entorno privado.

1.3. Objetivos del Trabajo

Los objetivos del trabajo son los siguientes:

- Adquisición de conocimiento del sector y de los dispositivos IOT.
- Lectura y comprensión de las técnicas del estado del arte en detección de anomalías en dispositivos conectados.
- Obtención de datos reales de conexiones a dispositivos IOT, en caso de no ser posible, generación de datos sintéticos.
- Prueba de las técnicas encontradas y evaluación en el problema actual.
- Investigación de algoritmos tradicionales de Machine Learning y su utilidad.
- Desarrollo de la solución utilizando los algoritmos más aptos.
- Evaluación de la detección de patrones y ataques de las técnicas utilizadas.
- Comparación de los resultados obtenidos con el estado del arte.

1.4. Descripción general del problema

La cantidad de dispositivos informáticos existentes que pueden ser víctimas de un ataque es masiva, por lo que securizar correctamente los dispositivos es fundamental. A pesar del uso de distintos protocolos de seguridad, se generan nuevos tipos de ataque todos los días, por lo tanto la detección es una necesidad a la hora de evitar los problemas derivados.

La pérdida de control de los dispositivos o la fuga de información de los mismos puede provocar pérdidas millonarias a las empresas o provocar una gran inseguridad a los usuarios de productos IOT, así como causar posibles daños a infraestructuras o personas. Por otro lado, muchos de los dispositivos pueden no tener la capacidad de computación necesaria para incluir una capa de seguridad robusta, que puede compensarse con una detección temprana.

1.5. Enfoque y método seguido

El enfoque pasa por conocer el problema y dar respuesta a preguntas específicas, mediante la aplicación de los conocimientos obtenidos y la evaluación de la propia aplicación de los mismos.

El método seguido durante el desarrollo se basa en el marco de trabajo *Agile* pensando en el desarrollo de la solución como un producto. De este modo se podrá iterar sobre una arquitectura definida y enfocarse en el desarrollo del producto en su totalidad, frente a un modelo tradicional de cascada donde cada fase de desarrollo recae en una sola parte del proyecto.

El formato de entrega será mediante un *minimum viable product* (MVP) en sprints de dos semanas durante el periodo de desarrollo y evaluación de las necesidades según la evolución del producto.

1.6. Planificación del Trabajo

Para la planificación del trabajo se han subdividido las tareas principales y se han mostrado en el siguiente diagrama de Gantt 1.1.

- Pec01 - Definición

- Pec01 - Planificación
- Pec02 - Búsqueda de fuentes del estado del arte
- Pec02 - Lectura de estado del arte
- Pec02 - Justificación de estado del arte
- Pec02 - Redacción
- Pec02 - Refinamiento de objetivos
- Pec03 - Planificación de Sprints
- Pec03 - Sprint 1
- Pec03 - Sprint 2
- Pec03 - Sprint 3
- Pec03 - Refinamiento/Redacción
- Pec04 - Revisión de apartados anteriores
- Pec04 - Redacción de nuevos apartados
- Pec05 - Presentación y defensa

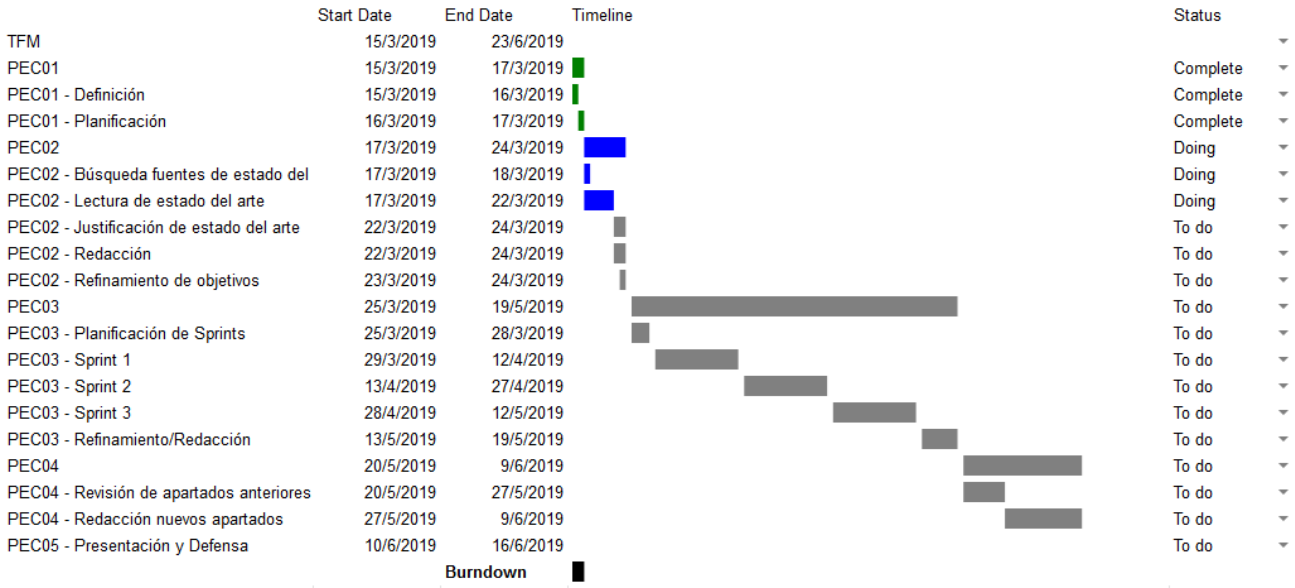

Figura 1.1: Planificación de tareas

Capítulo 2

Estado del Arte

Una de las aplicaciones más comunes dentro de la detección de anomalías es para la detección de intrusos (*Intrusion Detection*) y la creación de sistemas de detección de intrusos (*Intrusion Detection System, IDS*). Se trata de sistemas cuyo objetivo es la monitorización de los sistemas y redes informáticas, con el fin de alertar en caso de que puedan existir brechas de seguridad[1].

Con la ingente cantidad de redes y sistemas que se pueden encontrar a día de hoy es necesario incluir uno de estos sistemas, con el fin de mantener la integridad y disponibilidad de los mismos. Dentro de los IDS, se pueden identificar dos grandes implementaciones de estos sistemas, en los sistemas de detección de intrusos en Host (*Host-Based Intrusion Detection Systems, HIDS*) y los sistemas de detección de intrusos en red (*Network Intrusion Detection Systems, NIDS*).

- **HIDS:** Estos sistemas se caracterizan por implementarse en el Host utilizando información del propio sistema operativo para detectar actos maliciosos [2]. Esta información tiene distintos niveles de información, pero por lo general tienden a ser de bajo nivel sobre operaciones que se pueden estar realizando dentro del sistema. Esta información se consulta dentro de logs, por lo que el análisis de la información es más lento.

- **NIDS:** Para el segundo caso la monitorización se realiza a sistema de red, es decir, de comunicaciones entre distintos nodos y la monitorización de los paquetes que viajan entre ellos [1]. Esta información puede ser consumida en tiempo real, por lo que la reacción ante algún evento es más rápida que en los HIDS que necesitan revisar las acciones.

2.1. Métodos tradicionales de detección de anomalías

En la sección actual se describen algunos de los métodos tradicionales utilizados en IDS, tanto para HIDS como para NIDS.

- Network Security Monitor (NSM): se trata de uno de los primeros sistemas que permitió auditar el tráfico que circulaba dentro de la red [3]. El sistema escucha pasivamente dentro de la red y detecta si existe una conducta sospechosa al desviarse de patrones de conducta. La mayor parte de la monitorización se basa en protocolos estándar como *telnet, ftp, TCP/IP, etc.* por lo que le permitía utilizar una gran cantidad de datos heterogéneos.

- State transition analysis (USTAT): el sistema parte de que el host en un momento se encuentra en un estado seguro y que según las acciones que se realizan sobre el mismo el host cambia de estado, hasta que llega a un estado en el que compromete la seguridad [3]. Este sistema analiza los estados por los que ha pasado la máquina desde el estado seguro al comprometido.

- GrIDS: se trata de un IDS que utiliza un sistema de construcción de grafos basados en la red, donde cada nodo representa a un host y las aristas las conexiones entre los mismos. La representación gráfica de la actividad de la red permite ayudar al espectador en identificar qué está sucediendo [3].

- Haystack: en este caso el IDS se ayuda de métodos estadísticos para la detección de anomalías, definiendo estrategias para usuarios y grupos, además de definir variables del modelo como variables gaussianas independientes [4]. Para la detección se incluyen una serie de intervalos en los valores que en el momento que salen del rango normal, se calcula la distribución de probabilidades y si el *score* o puntuación es demasiado grande se genera una alerta.

Los métodos/sistemas listados se desarrollaron durante los años noventa, la tecnología ha evolucionado desde entonces y los sistemas se han vuelto más complejos y más propensos a los ciberataques. Por ello, se han desarrollado nuevas técnicas que se apoyan en el uso de técnicas de minería de datos (*Data Mining*) y las técnicas que se describirán a continuación de *Machine Learning*.

2.2. Machine Learning y la detección anomalías

El *Machine Learning* es una rama de la inteligencia artificial cuya premisa es hacer que la máquina aprenda una tarea sin haber sido específicamente programada para ello. El término fue descrito por Arthur L. Samuel en 1959 en un artículo en el que explica estudios de *Machine Learning* aplicado al juego de las damas [5], utilizando en una primera instancia métodos de aprendizajes más generales, como las redes neuronales de las que hablaremos más adelante, y otro métodos que tendrán que ser parametrizados para sus distintos usos.

La inteligencia artificial se puede entender como la inteligencia ejercida por las máquinas, al contrario que la inteligencia natural inherente a los humanos, que son capaces de realizar tareas cognitivas que permiten potenciar la resolución de las mismas e imitar comportamientos como el aprendizaje [6]. Dentro de la inteligencia artificial se pueden encontrar distintas definiciones según si se centran en el razonamiento o en el comportamiento:

- Actuar humanamente: este enfoque se basa en que las máquinas actúen como humanos más centrado en la interacción de máquinas con personas, más que en la resolución de problemas. Esta interacción se puede ver reflejada en el test ideado por Alan Turing en 1950, en el que se somete a la máquina ïnteligente.ª un interrogatorio realizado por un humano, donde éste no sabe que está hablando con una máquina, por lo tanto si es incapaz de detectar que se trata de una máquina esta ha actuado como un humano y puede considerarse inteligente.

- Pensar humanamente: esta vertiente se centra en imitar el pensamiento humano, entender cómo funcionan las mentes de los mismos y replicarlo en máquinas. Siguiendo esta corriente, si se consigue imitar el pensamiento humano, una máquina que resuelva el problema utilizará razonamientos humanos, en lugar de solucionar problemas a toda costa, independientemente de como lo realizan los humanos.

- Pensar racionalmente: basado en la lógica formal desarrollada en en siglos XIX y XX que permite formular los problemas en un lenguaje formal, utilizando el razonamiento matemático para resolver éstos.

- Actuar racionalmente: se centra en realizar el comportamiento más efectivo en un momento dado. Antes las distintas situaciones no siempre existe una acción correcta, pero si se puede llegar realizar una acción que minimice los riesgos.

Los enfoques mostrados conllevan sus propios enfoques filosóficos, sin embargo, han influenciado en gran medida a cómo se afrontan los problemas y cómo se han desarrollado las técnicas de inteligencia artificial que están en uso actualmente.

Como se ha comentado con anterioridad, el *Machine Learning* es una rama de la inteligencia artificial, dentro de la cual existe otra rama, *Deep Learning*. Las relaciones entre los términos se puede observar en la imagen 2.1 y como la inteligencia artificial engloba ambas ramas.

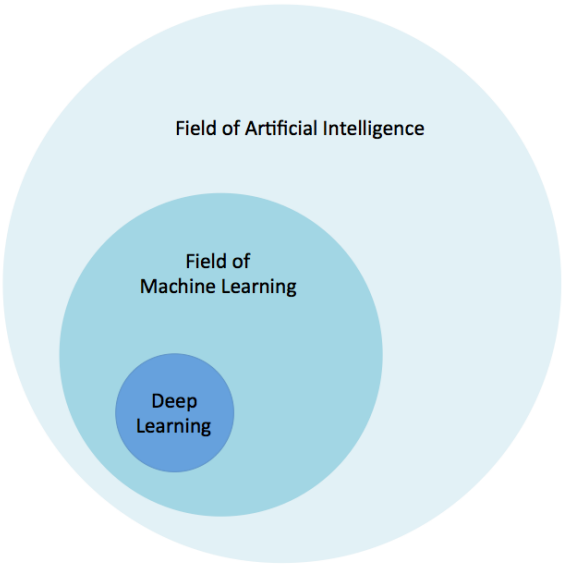

Figura 2.1: Relación entre Inteligencia Artificial, Machine Learning y Deep Learning

El *Deep Learning* se diferencia del *Machine Learning* convencional, en que están formados por modelos con distintas ramas de procesamiento que son capaces de extraer las representaciones de los datos con varios niveles de abstracción [7]. Mucha de la teoría relacionada se puede encontrar en los años 50 y 60, sin embargo, las aplicaciones y la capacidad de cómputo eran limitadas en la época, hasta ahora donde la capacidad de computación ha crecido a gran escala y en conjunto con las nuevas técnicas propuestas, una mayor cantidad de datos y la aplicación de transformaciones no lineales hacen que esta vertiente viva su época dorada.

Una de las implementaciones más representativas del *Deep Learning* se trata de las redes neuronales, las cuales se encuentran basadas en el funcionamiento del cerebro humano, utilizando el concepto de neuronas que pueden realizar cálculos, la conexión entre las mismas, la función de realizar una tarea específica, etc. Añadir, que se encuentra basado y que el cerebro humano es un sistema mucho más complejo que las redes neuronales y tiene muchos más comportamientos [8].

Bibliografía

[1] Rebecca Bace and Peter Mell. Intrusion detection systems. *National Institute of Standards and Technology (NIST)*, 2001.

[2] Giovanni Vigna and Christopher Kruegel. Host-based intrusion detection. 2005.

[3] Stefan Axelsson. Research in intrusion-detection systems: a survey. *Department of Computer Engineering, Chalmers University of Technology*, 1998.

[4] Pedro Garcia-Teodoro, Jesus Diaz-Verdejo, Gabriel Maciá-Fernández, and Enrique Vázquez. Anomaly-based network intrusion detection: Techniques, systems and challenges. *computers & security*, 28(1-2):18–28, 2009.

[5] Arthur L Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 44(1.2):206–226, 2000.

[6] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.

[7] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.

[8] Simon Haykin. *Neural networks*, volume 2. Prentice hall New York, 1994.