Universidad del Valle de Guatemala
Facultad de Ingeniería
Departamento de Ciencias de la Computación
CC3094 - Security Data Science

Proyecto Final - Fase #1

Evelyn Andrea Amaya Malin - 19357
Brandon Josue Hernández Marroquín - 19376
Oscar André Paredez Urizar - 19109
Guido Sebastián Padilla Aldana - 19200

Guatemala, Ciudad de Guatemala 21 de febrero de 2023

Problema

Modelo de detección de patrones maliciosos dentro de una red por medio de movimientos gestionados como peticiones, accesos, etc los cuales ponen en peligro la integridad y bienestar de la información que fluye dentro de esa misma red.

Motivación

El objetivo del proyecto es entrenar un modelo para que detecte patrones dentro de la red haciendo uso de la actividad (logs) registrados en la nube y así identificar posibles ataques. El propósito es mantener segura la información de los usuarios conectados a la red, alertar a los administradores en caso de encontrar una actividad sospechosa y bloquear al usuario detectado como malicioso. De esta forma, al estar conectado a una red, el usuario no se tendrá que preocupar por la seguridad de su información como credenciales de bancas en línea, correos o mensajes confidenciales, información personal, etc.

Alcance

Monitorear el estado de la red de forma recurrente para detectar patrones de posibles ataques y bloquearlos antes de que ponga en riesgo los datos en tránsito de una red.

Objetivos

- Detectar movimientos sospechosos dentro de la red de una manera sencilla y valiosa.
- Mejorar el tiempo de respuesta de ataques tras detectar patrones de uso de la red.
- Notificar o alertar posibles ataques de la red para que sean tratados por un administrador.
- Bloquear usuarios con actividad sospechosa del uso de la red.

Preguntas clave

- ¿Qué variables son las más importantes al momento de detectar a un intruso dentro de la red?
- ¿Qué servicios o recursos de la información son más propensos a ser más accedidos durante un ataque?
- ¿Será prudente tener tres modelos diferentes para poder reducir los errores de falso negativo [1]?

[1] Se identifica como falso negativo un ataque malicioso que el modelo marque como inofensivo.

Revisión de literatura

- **Android malware detection using network traffic based on sequential deep learning models**

Esta referencia es útil debido a que es un antecedente que valida la idea de encontrar patrones maliciosos dentro de la red:

Este paper se enfoca en la detección de malware insertado a dispositivos móviles Android mientras esté produce tráfico dentro de una red. En este caso combinan un dataset público y uno privado generado por ellos mismos, con el fin de cubrir todas las alternativas posibles al momento de la detección. Se realizan detecciones de anomalías mediante la información básica del tráfico generado, el momento, el tipo de conexión y el tipo de contenido.

- **Detection of Man In The Middle Attacks in Wi-Fi networks by IP Spoofing**

Esta referencia es útil debido a que es un antecedente que valida la idea de encontrar patrones haciendo uso de la actividad dentro de la red:

En un ataque Man-in-the-Middle el atacante es un intermediario entre el router y el cliente. El atacante es capaz de monitorear todo el tráfico y afectar la confidencialidad e integridad de los mensajes del cliente. Se describe cómo se puede usar la herramienta airmon-ng para monitorear todo el tráfico de una red wifi y cómo al usar wireshark se puede obtener la dirección MAC del usuario a ser atacado. Tras tener acceso a la red y a la dirección MAC del usuario, el atacante puede ver ahora cada petición que realiza el cliente y luego mandarla al router sin ninguna modificación. El router realiza un experimento sobre configurar una segunda red para luego comparar los tiempos de llegada de la información del cliente contra la primera red. En dado caso que el tiempo sea mayor, significa que hay un MIM.

- **Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things**

Esta referencia es útil debido a que es un antecedente que valida la idea de usar un modelo de Machine Learning luego de habernos enfocado en un área específica de la ciberseguridad:

Este paper presenta el análisis y enfoque que se puede dar con respecto a los modelos de Machine Learning para prevenir ciberataques. Se detalla cómo el campo es enorme y como este estudio se decidió enfocar en específicos protocolos, haciendo énfasis en backdoors, SQL injections y command injections. Además, explica sobre el internet de las cosas y todos los riesgos a los que está expuesto el usuario tomando en cuenta la posición del programador. Por ejemplo, se habla sobre la vulnerabilidad debido a los accesos sin credenciales a ciertos sistemas o ingreso de credenciales a una plataforma sin límite de intentos lo cual puede hacer uso de fuerza bruta.

- **Machine Learning aplicado en Sistemas de Detección de Intrusos**

Esta referencia es útil debido a que es un antecedente que presenta a los Sistemas de Detección de Intrusos y porque se le debería de aplicar ML, además de brindar teoría fundamental para el proyecto:

En este paper se presenta cómo es que se utilizan los Sistemas de Detección de Intrusos (SDI) para detectar a los intrusos en una red con base en logs. Menciona cómo es que estos sistemas se mantienen mediante configuraciones estáticas y el administrador debe de velar por la actualización de las mismas para que no se vea comprometida la información valiosa. Por lo que se ha intentado implementar ML con diferentes modelos como lo son Árboles de decisión, Reglas de decisión, Redes neuronales, Algoritmos genéticos entre otros. Además de brindar esquemas de recolección de datos, que pueden ser por medio del host que sea el SDI y por medio del tráfico de la red. Dando un punto de vista del estado del arte para este tema.

-

Recolección de datos inicial

Los datos iniciales se planean buscar en proyectos que sean de un ámbito similar al del proyecto actual. Estos datasets los brindan diferentes páginas en competencias que hacen, por ejemplo Kaggle es una página en donde se realizan este tipo de competiciones y brindan el dataset. De no encontrar la data online y de carácter público, se buscará alguna empresa que haya sido vulnerada y si luego de esto no se puede encontrar el grupo generará los datos necesarios con algún distribuidor de nube y los logs.

En este este caso se encontró en la fuente de Kaggle un dataset con las siguientes características, los cuales no cuentan con información delicada ya que son previamente filtrados por esta compañia: duration, protocol_type, service, flag, src_bytes, dst_bytes, land, wrong_fragment, urgent, hot, num_failed_logins, logged_in, num_compromised, root_shell, su_attempted, num_root, num_file_creations, num_shell, ,num_access_files, num_outbound_cmds, is_host_login, is_guest_login, count, srv_count, serror_rate, srv_serror_rate, rerror_rate, srv_rerror_rate, same_srv_rate, diff_srv_rate, srv_diff_host_rate, dst_host_count, dst_host_srv_count, dst_host_same_srv_rate, dst_host_diff_srv_rate, dst_host_same_src_port_rate, dst_host_srv_diff_host_rate, dst_host_serror_rate, dst_host_srv_serror_rate, dst_host_rerror_rate, dst_host_srv_rerror_rate, attack y level.

https://www.kaggle.com/datasets/hassan06/nslkdd

Referencias Bibliográficas

Anand, Gokul & Prathiba, Sahaya Beni & Gunasekaran, & Ponmani,. (2018). Detection of Man In The Middle Attacks in Wi-Fi networks by IP Spoofing. 319-322. 10.1109/ICoAC44903.2018.8939063.

Fallah, S, Bidgoly, AJ. (2022) *Android malware detection using network traffic based on sequential deep learning models*. Softw Pract Exper. 52(9): 1987–2004. doi:10.1002/spe.3112

M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan and R. Jain. (Aug. 2019). *Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things*. IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6822-6834., doi: 10.1109/JIOT.2019.2912022.

I. Peluffo, M. Capobianco & J. Echaiz. (2014). *Machine Learning aplicado en Sistemas de Detección de Intrusos*.