Detección de ciberataques mediante aprendizaje profundo

Christian Garzón
GITA Lab
Faculty of Engineering
University of Antioquia UdeA
Medellín, Colombia
christianc.garzon@udea.edu.co

***Index Terms*—Software Defined Networks (SDN), P4 language, Programmable Data Plane, OpenFlow, Machine Learning, Firewall, IDS, IPS, DDoS, Spoofing.**

I. CONTEXTO

La seguridad es un factor esencial en cada actividad humana, especialmente en lo que se refiere a la información privada. Un activo de información puede encontrarse digitalmente, cambiando al concepto de ciberseguridad. Además, esta información pertenece a un individuo o a una organización, y parte de ella debe permanecer confidencial. Según Kaspersky [1], la ciberseguridad es la práctica de defender los dispositivos informáticos como ordenadores, servidores, teléfonos móviles, dispositivos móviles, redes y datos. Sin embargo, ¿de qué se defiende la ciberseguridad? Las ciberamenazas se han ido desarrollando desde los orígenes del ordenador; hoy en día, se dirigen a las redes, las aplicaciones, la información y las operaciones. Como explica Kaspersky, en el ámbito de la ciberseguridad se desarrollan y utilizan a diario muchos ciberataques; los actores están motivados por el dinero, la política, el terrorismo y otros [1].

La seguridad de la información ha pasado a ser un objetivo inestimable para cualquier organización, aunque implique un aumento de los precios. Actualmente, los sistemas de comunicación tradicionales poseen ciertas vulnerabilidades, las cuales podrían ser usadas maliciosamente por un tercero; ahora bien, los dispositivos de ciberseguridad, también tradicionales, podrían aumentar los retrasos adicionales y agregar un tráfico duplicado a la red. Lo anterior, es una situación que no puede ser escalable en determinadas infraestructuras.

Hoy en día, el IoT (en inglés Internet of Things) y aplicaciones industriales utilizan las tecnologías de comunicación tradicionales (basadas en TCP/IP o Ethernet) para supervisar y controlar sus sistemas. Evidentemente, esto hace que hereden indirectamente vulnerabilidades críticas de ciberseguridad. Sin embargo, nuevas técnicas y modernas herramientas, que se enfocan en la detección y mitigación de ciberataques mediante la clasificación de los flujos de red, incluyen el uso de algoritmos de machine learning, los cuales podrían alcanzar un mejor rendimiento, incluso a menor costo. Hoy en día, los ataques más sofisticados y peligrosos pueden aprovechar las brechas en las redes, y cuanto más se hayan desarrollado, mejores sistemas de defensa deberá utilizar el objetivo atacado. Los principales ataques son:

- MALWARE.
  - Virus.
  - Troyanos.
  - Spyware.
  - Ransomware.
  - Botnets.
- PHISHING.
- MAN-IN-THE-MIDDLE.
- DENIAL-OF-SERVICE.

II. OBJETIVO DE MACHINE LEARNING

Para realizar la clasificación que indique si se está bajo un ataque o no (clasificación binaria), es necesario tener un contexto de los flujos de red previamente analizados. Los mencionados flujos de red pueden verse como una serie de entradas que generan diversas salidas, es decir una secuencia de flujos. Por lo anterior, se hace necesario el uso de técnicas de modelamiento secuencial.

III. DATASET

El dataset a usar se puede encontrar en [2], este contiene una serie de datos benignos y malignos (flujos de ataques) que son procesados con CICFlowMeter a partir de estampas de tiempo, IP fuente y destino, puertos fuente y destino, protocolo y ataque; de estos se genera una cantidad de 83 características. Este dataset está etiquetado según el ataque. Los ciberataques usados en este dataset fueron Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet y DDoS. El tamaño del dataset varia según el día de prueba, lunes a viernes, para un total de unas 51.1 Gb con distribución L:11, M:11, W:13, J:7.8 y V:8.3. Las etiquetas y la cantidad de elementos se muestran en la figura 1.

IV. METRICAS DE DESEMPEÑO

Tratándose de un problema de clasificación podrá usarse una matriz de confusión de dos clases que poseerá.

Class Labels	Number of instances
BENIGN	2359087
DoS Hulk	231072
PortScan	158930
DDoS	41835
DoS GoldenEye	10293
FTP-Patator	7938
SSH-Patator	5897
DoS slowloris	5796
DoS Slowhttptest	5499
Bot	1966
Web Attack – Brute Force	1507
Web Attack – XSS	652

Fig. 1. Clases de etiquetas en dataset CICIDs2017 [2].

TABLE I
MATRIZ DE CONFUSIÓN

	EXPECTED	
PRE	True Positives TP	False Positives FP
	False Negatives FN	True Negatives TN

A. PRECISION

Responde que número de positivos fueron reportados e identificados de forma correcta.

$$\frac{TP}{TP + FP} \tag{1}$$

B. RECALL

Indica el porcentaje de positivos reales que se identificaron. Para este caso, la diferencia con (1) radica en que (2) tiene en cuenta los falsos negativos.

$$\frac{TP}{TP + FN} \tag{2}$$

C. F1 SCORE

Medida harmonica que relaciona (1) y (2).

$$\frac{2 * TP}{2 * TP + FP + FN} \tag{3}$$

V. REFERENCIAS Y RESULTADOS PREVIOS

Algunos autores ya han hecho uso de algoritmos como RNN (Recurrent Neural Network) y obtenido interesantes resultados. En [3] se obtuvieron resultados con un 0.998 y 0.919 en el F-Score para detección de incidentes y fraudes respectivamente. Por otra parte, en [4] haciendo uso de un dataset y con entradas solo basadas en IP destino y fuente alcanzó una detección del 0.912.

REFERENCES

[1] What is Cyber Security?. Kaspersky. [Online]. Available: https://www.kaspersky.com/resource-center/definitions/what-is-cyber-security.

[2] Panigrahi, Ranjit Borah, Samarjeet. (2018). A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems. International Journal of Engineering Technology. 7. 479-482.

[3] KP, Soman, et al. RNNSecureNet: Recurrent neural networks for Cyber security use-cases. arXiv preprint arXiv:1901.04281, 2019.

[4] BEN FREDJ, Ouissem, et al. CyberSecurity attack prediction: a deep learning approach. En 13th International Conference on Security of Information and Networks. 2020. p. 1-6.