Contexto de aplicación.

Entre los ciberataques más comunes se encuentran la "denegación de servicios distribuida" (DDoS o Distributed Denial of Service). Una DDoS ocurre cuando los atacantes utilizan múltiples dispositivos para agobiar un sistema, red o host objetivo y disminuir sus recursos, lo cual, abruma la capacidad del objetivo para manejar solicitudes legítimas, volviéndola inaccesible para los usuarios legítimos. Los informes de Cisco-talos [1] y la división europea del CERT [2] vinculan el incremento de este tipo de ataques a las guerras entre Ucrania-Rusia, Israel-Hamás y el interés de terceros países en las regiones, donde se presenta el conflicto. Además de las referencias citadas, para otras estadísticas de ciberataques, se puede consultar [3]

Para contra-restar un DDoS, el primer paso es identificarlo y una opción es a través de un sistema de detección de intrusiones (IDS o Intrusion Detection System), el cual, se debe caracterizar por su velocidad para detectar irregularidades, precisión y confiabilidad. La operación de un IDS consiste en una clasificación binaria del tráfico bajo análisis (normal o anómalo) o múlticlase (además, del tráfico normal, se identifican varios escenarios: R2R, U2R, DoS, DDoS, Probe, entre otros). Entre las técnicas de detección de intrusión, tenemos:

- Detección anómala: Se verifica la actividad del sistema y se clasifica el tipo de tráfico (normal o anómalo). Generalmente, consiste de dos fases: 1. Entrenamiento: Se determina el perfil de conducta normal. 2. Prueba: El tráfico actual se compara con el perfil obtenido en la fase 1, para su identificación.
- Técnica de detección de uso indebido: Se detectan ataques a través de perfiles pre-configurados (patrones, firmas, otros); en este caso, se "conoce" la conducta anormal del sistema y si el tráfico no la cumple, se considera normal.

Para ampliar el tema de IDS, consultar [4], [5] y [6].

La evolución y crecimiento de Internet generó nuevos paradigmas para diseñar, implementar y gestionar redes de telecomunicaciones con características tales como, escalabilidad, seguridad, flexibilidad, desempeño, entre otras. Dentro de los paradigmas propuestos se encuentran las redes programables y uno de sus enfoques corresponde a las redes definidas por software (SDN o Software Defined Networks). La principal propiedad de la arquitectura SDN es la habilidad para separar las funciones de envió y control de la red, lo cual: 1. Facilita al administrador la gestión de dispositivos en la red, sin importar el vendedor. 2. Reduce la intervención humana.

La arquitectura lógica de SDN esta normalizada por la ONF y se describe como:

- Plano de aplicación: se implementa a través de APIs, las cuale sexpandes las capacidades de los servicios SDN.
- Plano de control: Se implementa a través de controladores; los ontroladores son los responsables de informar las reglas de flujo en la red.
- Plano de datos: Se implementa a través de los dispositivos de interconexión y host, los cuales cumplen las reglas informadas por el controlador.

La comunicación entre el plano de control y de datos, se da a través del protocolo Openflow. Para ampliar el tema de SDN, consultar [7], [8] y [9]. En una red implementando arquitectura SDN, los posibles objetivos son: 1. Ataque al plano de datos (ataque en SDN y redes tradicionales). 2. Ataque a Openflow (ataque en SDN). 3. Ataque al plano de control (ataque en SDN). 4. Ataque al plano de aplicación (ataque en SDN y redes tradicionales). Para cualquiera de los cuatro objetivos señalados, el DDoS es una opción de ataque.

Objetivo de machine learning (queremos predecir X, dada tal información)

El objetivo del presente proyecto es predecir, a través de Redes Neuronales Convolucionales (CNN o Convolutional Neural Networks) convolucional CNN si el tráfico recolectado en el dataset InSDN (84 clases y 343.939 registros para cada clase) es anómalo y en caso afirmativo clasificar el tipo de ataque (DDoS, Probe y otros).

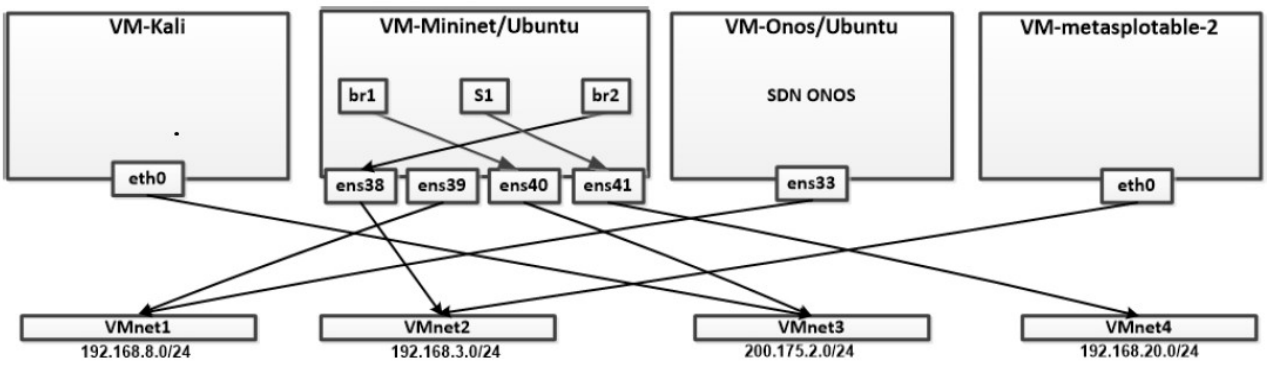
Dataset: tipo de datos, tamaño (número de datos y tamaño en disco), distribución de las clases

En el caso de la seguridad informática, un problema es la falta de dataset para el tráfico de red y la detección de intrusiones (debido a la privacidad y las preocupaciones legales); además, en el caso de SDN se utilizan datos originados en redes convencionales, pero ¿afectará esto la predicción?.

InSDN [10] es un dataset obtenido al implementar SDN e incluye tráfico normal (HTTPS, HTTP, SSL, DNS, correo electrónico, FTP, SSH) y diferentes tipos de ataques(DoS, DDoS, brute force attack, web applications, exploitation, probe, and botnet). Información general del dataset:

Grupo	Distribución del tráfico	Número de instancias	Total	Tamaño PCAP
Normal	Skype, Facebook, File Transfer, Youtube, Email, DNS, Chat, Browsing	68424	68424 (19.90%)	3.58 GB
Metasploitable-2	DDoS	73529	136743 (39.76%)	669 MB
	Probe	61757		
	DoS	1145		
	brute-force-attack	295		
	Exploitation (R2L)	17		
OVS	DoS	52471	138772 (40.34%)	1.21 GB
	DDoS	48413		
	Probe	36372		
	brute-force-attack	1110		
	Web_attack	192		
	Botnet	164		

Originalmente, el dataset se dividió en tres grupos: 1. Normal: Incluye únicamente el tráfico normal. 2. Mealsplotable-2: Contiene los tráficos de ataque dirigidos al servidor mealsplotable 2. 3. OVS: Consideramos los ataques a la máquina OVS. El conjunto de datos se generó utilizando cuatro máquinas virtuales (1. Kali Linux y representa el servidor atacante. 2. Ubuntu 16.4 y actúa sobre el controlador ONOS. 3. Ubuntu 16.4 que servirá para el conmutador Mininet y OVS. 4. Linux basada en Metaploitable 2 para proporcionar servicios vulnerables) y además, se crearon cuatro hosts virtuales (Vhost) utilizando la herramienta Mininet (los dos primeros Vhosts (h1 y h2) generan tráfico malicioso, mientras h3 representa las actividades normales y h4 actúa como un servidor web). Testbed:



Se representaron escenarios con la fuente de ataque ubicada tanto del exterior, como del interior de la red SDN. El tráfico de datos se capturó utilizando la herramienta Wireshark y las capturas se guardaron en formato PCAP.

Métricas de desempeño (de machine learning y negocio)

Se han utilizado diferentes indicadores de desempeño para evaluar la eficiencia de las técnicas de aprendizaje supervisado, incluyendo a: F-score, recall and precision.

En el caso de redes en operación es de interés, el desempeño del controlador SDN y a nivel de la red, caracterísiticas como el delay y el throughput.

Resultados previos

Ejemplos de la predicción de tráfico con CNN para el dataset InSDN, se pueden hallar en: [11], [12], [13], [14], [15], [16], [17].

Referencias

[1] https://blog.talosintelligence.com/cisco-talos-2023-year-in-review/

[2] https://www.ccn-cert.cni.es/es/informes/informes-ccn-cert-publicos/7188-ccn-cert-ia-35-23-ciberamenazas-y-tendencias-edicion-2023/file.html

[3] https://www.forbes.com/advisor/education/it-and-tech/cybersecurity-statistics/

[4] A. K. Saxena, S. Sinha and P. Shukla, "General study of intrusion detection system and survey of agent based intrusion detection system," 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 2017, pp. 471-421, doi: 10.1109/CCAA.2017.8229866.

[5] A. R. b. Gupta and J. Agrawal, "A Comprehensive Survey on Various Machine Learning Methods used for Intrusion Detection System," 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), Gwalior, India, 2020, pp. 282-289, doi: 10.1109/CSNT48778.2020.9115764.

[6] A. L. Giri and S. Annamalai, "Intrusion Detection System for Local Networks – A Review Study," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 1388-1393, doi: 10.1109/ICACITE53722.2022.9823433.

[7] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," in Proceedings of the IEEE, vol. 103, no. 1, pp. 14-76, Jan. 2015, doi: 10.1109/JPROC.2014.2371999.

[8] B. A. A. Nunes, M. Mendonca, X. -N. Nguyen, K. Obraczka and T. Turletti, "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," in IEEE Communications Surveys & Tutorials, vol. 16, no. 3, pp. 1617-1634, Third Quarter 2014, doi: 10.1109/SURV.2014.012214.00180.

[9] J. Xie et al., "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," in IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 393-430, Firstquarter 2019, doi: 10.1109/COMST.2018.2866942.

[10] M. S. Elsayed, N. -A. Le-Khac and A. D. Jurcut, "InSDN: A Novel SDN Intrusion Dataset," in

IEEE Access, vol. 8, pp. 165263-165284, 2020, doi: 10.1109/ACCESS.2020.3022633

[11] H. A. Hassan, E. E. Hemdan, W. El-Shafai, M. Shokair and F. E. A. El-Samie, "An Efficient Intrusion Detection System for SDN using Convolutional Neural Network," 2021 International Conference on Electronic Engineering (ICEEM), Menouf, Egypt, 2021, pp. 1-5, doi: 10.1109/ICEEM52022.2021.9480383.

[12] A. H. Janabi, T. Kanakis and M. Johnson, "Convolutional Neural Network Based Algorithm for Early Warning Proactive System Security in Software Defined Networks," in IEEE Access, vol. 10, pp. 14301-14310, 2022, doi: 10.1109/ACCESS.2022.3148134.

[13] M. K, K. K, U. J, P. T, S. AR and S. D, "Classification of DDoS in Software-Defined IoT Networks using Hybrid Feature Selection Technique with DL," 2022 Third International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 2022, pp. 1-6, doi: 10.1109/ICSTCEE56972.2022.10099827.

[14] Y. Wang, X. Yi, Z. Bao and B. Yu, "An Effective Anomaly Detection Model Combined Feature Selection with Improved CNN in SDN," 2023 3rd Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), Shenyang, China, 2023, pp. 281-287, doi: 10.1109/ACCTCS58815.2023.00048.

[15] H. A. Hassan, E. El-Din Hemdan, M. Shokair, F. E. A. El-Samie and W. El-Shafai, "An Efficient Attack Detection Framework in Software-Defined Networking using Intelligent Techniques," 2023 3rd International Conference on Electronic Engineering (ICEEM), Menouf, Egypt, 2023, pp. 1-6, doi: 10.1109/ICEEM58740.2023.10319575.

[16] R. B. Said and I. Askerzade, "Attention-Based CNN-BiLSTM Deep Learning Approach for Network Intrusion Detection System in Software Defined Networks," 2023 5th International Conference on Problems of Cybernetics and Informatics (PCI), Baku, Azerbaijan, 2023, pp. 1-5, doi: 10.1109/PCI60110.2023.10325985.

[17] R. Ben Said, Z. Sabir and I. Askerzade, "CNN-BiLSTM: A Hybrid Deep Learning Approach for Network Intrusion Detection System in Software-Defined Networking With Hybrid Feature Selection," in IEEE Access, vol. 11, pp. 138732-138747, 2023, doi: 10.1109/ACCESS.2023.3340142