Detección de DDoS con Deep CNN-LSTM: un análisis de calibración, umbral operativo y latencia

Autora

Joanna Carrión

Mentor

PhD. Yezid Donoso

Contactos

joannacarrion14@gmail.com

https://www.linkedin.com/in/joanna-carrion-perez/

https://orcid.org/0000-0002-8967-1278

ydonoso@uniandes.edu.co

Reseña profesional

Bachiller en Ingeniería Electrónica con formación en IA y sistemas embebidos. Actualmente, trabaja en OSINERGMIN apoyando en la supervisión del sector eléctrico. Fue becaria ELAP y realizó una pasantía de investigación en la University of Alberta en IA aplicada. Intereses en análisis de datos, con enfoque multidisciplinario e innovador.

Reconocimientos

Agradezco a mi mentor por su guía técnica y acompañamiento en cada etapa del proyecto, al **Programa de Mentoreo IT Women de LACNIC** por la formación, el seguimiento y las oportunidades brindadas.

También a mi familia por su apoyo constante, paciencia y motivación que hicieron posible este trabajo.

Resumen del trabajo técnico

Se presenta un sistema de detección de DDoS basado en un modelo Deep CNN-LSTM entrenado de forma causal con ventanas 32×10 y validación rigurosa sin fuga (escalado ajustado sólo con entrenamiento). La arquitectura combina convoluciones 1D dilatadas con bloques residuales, una capa LSTM y atención temporal para capturar patrones multiescala y la dinámica de tráfico. El trabajo operacionaliza el clasificador: se calibran las salidas mediante temperature scaling, se fija un umbral según una política de error, y se implementa inferencia en streaming con medición de latencia y throughput. Se evalúa con curvas ROC/PR, matriz de confusión y métricas por clase, y se analiza la robustez frente a ruido y pérdida parcial de variables. El pipeline es reproducible en entorno Colab y se incluye un stub de mitigación (bloqueo) para futura integración. En conjunto, se demuestra un detector consistente, con bajo falso positivo bajo la política elegida y alta recuperación de ataques, listo para ser extendido a escenarios y despliegues reales.

Introducción: antecedentes, descripción del trabajo y resultados esperados

Los ataques DDoS siguen en aumento y, aunque muchos trabajos con redes profundas reportan alta precisión, suelen carecer de evaluación causal, control de fuga, calibración de probabilidades y métricas de latencia para decisión en línea [1] [2].

Se diseña un pipeline reproducible para detección de DDoS con un modelo CNN-LSTM robusto que combina convoluciones 1D dilatadas, bloques residuales, una capa LSTM y atención temporal. El entrenamiento es causal con ventanas de 32×10; la partición es sin fuga (escalado ajustado solo con *train*). La validación usa AUC-PR con parada temprana y reducción de la tasa de aprendizaje. Se calibra con temperature scaling y se fija un umbral operativo guiado por FPR objetivo. Se implementa streaming y se miden latencia y robustez (ruido y pérdida de variables). La mitigación queda como guión simulado.

Los resultados esperados son: PR-AUC y ROC-AUC ≥0.98, exactitud ≥0.99, FPR ≤1% con recall ≥0.98, latencia ≤150 ms y degradación ≤5 pp.

Metodología

Se usó el dataset APA-DDoS [3] respetando el orden temporal. Se realizó partición 70/15/15 sin fuga (escalado solo con entrenamiento) y se generaron ventanas causales 32×10 etiquetadas por el último paso. El modelo es un CNN-LSTM robusto con convoluciones dilatadas y bloques residuales, seguido de LSTM y atención temporal. Se entrenó con Adam y entropía cruzada binaria, con suavizado de etiquetas, L2 y pesos por clase. La validación empleó AUC-PR, parada temprana y reducción de tasa de aprendizaje. Se calibró con temperature scaling, se fijó el umbral por FPR objetivo y se evaluó en streaming, latencia y robustez; mitigación en stub.

Gráficos, datos y principales resultados

Con calibración T=1.117 y umbral operativo θ=0.9644 (FPR ≤ 1 %); como se aprecia en la Fig. 1, el conjunto de prueba muestra TN=341, FP=1, FN=0 y TP=366. Esto implica accuracy=0.9986, recall de ataques 1.000, precisión=0.9973, F1=0.9986 y FPR≈0.29 % (1/342). En la Fig. 2, se observa AP=0.999863 y un tramo casi plano de precisión hasta recall cercano a 1.0, lo que confirma alta separabilidad. Adicionalmente, las curvas globales reportan PR-AUC de 0.999863 y ROC-AUC de 0.999856. La distribución de puntuaciones sitúa 366 ataques por encima del umbral y 341 benignos por debajo, con 1 benigno por encima. En latencia, se midió media 116.877 ms por ventana, p95=152.027 ms, máximo 448.915 ms y throughput aproximado 8.6 ventanas por segundo, respaldando bajo falso positivo y recuperación total.

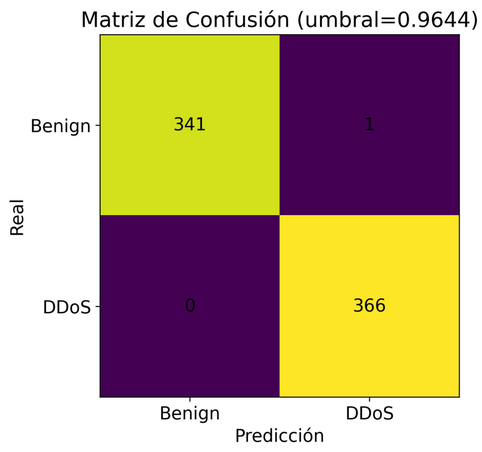

Fig. 1. Matriz de confusión

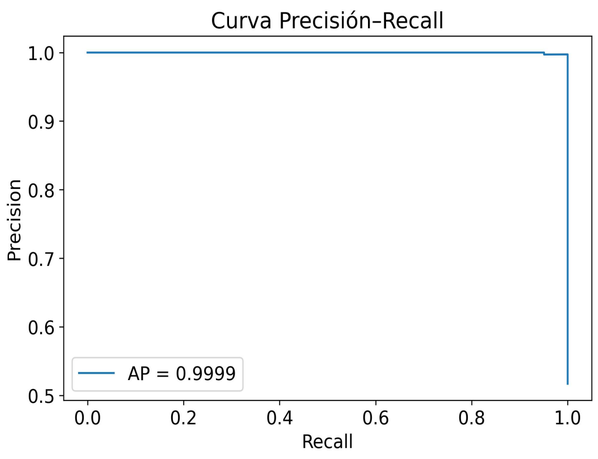

Fig. 2. Gráfica de la curva precisión-recall

Relevancia del trabajo técnico

El proyecto convierte un clasificador profundo en un sistema operativo: entrenamiento causal sin fuga de datos, validación con AUC-PR, calibración de probabilidades y umbral definido por una política de error (objetivo sobre FPR). La arquitectura CNN-LSTM con convoluciones dilatadas, bloques residuales y atención temporal capta patrones multiescala y concentra la evidencia más útil, elevando robustez e interpretabilidad. Además, la evaluación incluye streaming y latencia con lote unitario, de modo que las métricas se conectan con requisitos de operación en tiempo casi real.

El pipeline es reproducible en Colab, modular y listo para extender: admite nuevos datasets, ajuste del umbral por objetivos distintos, y exportación del modelo para despliegue posterior. Se documenta robustez ante ruido y pérdida de variables, y se deja un stub de mitigación que permite integrar acciones de bloqueo en un entorno real o SDN. En conjunto, el trabajo cierra la brecha entre resultados académicos y decisiones prácticas.

Conclusiones, desafíos y oportunidades

El sistema de detección basado en Deep CNN-LSTM, entrenado de forma causal y sin fuga, demostró operación práctica mediante calibración por temperatura y un umbral guiado por una política de bajo falso positivo. Con θ=0.9644 y T=1.117, se alcanzó recuperación total de ataques; FPR≈0.29%, exactitud 0.9986 y métricas cercanas a uno, con latencia media de 117 ms por ventana en streaming.

Sin embargo, persisten ciertos desafíos como validar con otros conjuntos y capturas reales, manejar deriva y variaciones de tráfico, ampliar pruebas de robustez ante pérdida de variables, y caracterizar el rendimiento bajo cargas concurrentes y tamaños de ventana. También se debe consolidar la estimación de incertidumbre y estabilidad de la calibración en el tiempo.

Por otro lado, las oportunidades incluyen integrar mitigación real mediante listas de bloqueo o SDN, extender a multiclase, optimizar throughput con lotes y exportar a TFLite u ONNX y ajustar el umbral en línea.

Referencias

[1] G. Ariel, “La revolución silenciosa del IoT: conectar el mundo, transformar vidas,” Telefónica. [En línea]. Disponible en: https://www.telefonica.com/es/sala-comunicacion/blog/revolucion-silenciosa-iot-conectar-mundo-transformar-vidas/

[2] Radware, “What is the Mirai Botnet?” [En línea]. Disponible en: https://www.radware.com/security/ddos-knowledge-center/ddospedia/mirai/

[3] Y. Reddy, “APA-DDoS Dataset.” [En línea]. Disponible en: https://www.kaggle.com/datasets/yashwanthkumbam/apaddos-dataset/data

Citación de esta publicación

J. Carrión. “Detección de DDoS con Deep CNN-LSTM: un análisis de calibración, umbral operativo y latencia” [Presentación de póster]. Presentado en: LACNIC 44-LACNOG 2025, 6-10 de octubre de 2025, San Salvador, El Salvador.