Machine Learning as a Defense Against Distributed Denial-of-Service Attacks

Carl Lochstampfor Jr

Department of Cybersecurity, Old Dominion University

CYSE 420 — Applied Machine Learning in Cybersecurity

August 29, 2026

Machine Learning as a Defense Against Distributed Denial-of-Service Attacks

Exploring ML Benefits

Early and traditional defense systems check network traffic against known signatures, allowing anything new or uncatalogued to bypass those static rules (e.g., zero-day attacks). Machine learning takes a different approach by building a statistical baseline of normal behavior, such as packet rates, protocol mix, and source address diversity, and then scoring deviations. This matters in practice: a model can spot credential stuffing across thousands of addresses or lateral movement that appears normal on a single host but stands out when viewed across the whole network.

Mitigating Cyber Threats: DDoS

Detecting distributed denial-of-service attacks works well with supervised learning. Flow records provide metrics such as packets per second, source IP entropy, and the ratio of SYN to ACK packets. A random forest trained on labeled examples of normal and attack traffic can classify flows in milliseconds. Najafimehr et al. (2023) also point out that unsupervised clustering helps identify novel attack vectors that carry no labels.

Challenges and Considerations

Malicious traffic is far less common than normal traffic, so imbalanced datasets push models toward predicting "benign," and labeled network data is hard to obtain in the first place. Vassilev et al. (2025) show that evasion and data poisoning can occur at every stage of the ML lifecycle, which makes the detector itself an attack target. False positives are costly here, since dropping a legitimate flow causes the outage the model was meant to prevent.

Real-World Applications

Cloudflare (2026) reported that it stopped 935 network-layer attacks at over 1 terabit per second in the first half of 2026. Most attacks, about 90.60%, ended within ten minutes, and some record-breaking bursts lasted just 35 seconds. At that speed, alerts reach human analysts after the attack is over, underscoring why automated classification has replaced manual review.

References

Cloudflare. (2026, August 11). Cloudflare DDoS threat report H1 2026: 1 Tbps attacks soar as DNS floods and geopolitical tensions drive a new wave. https://blog.cloudflare.com/ddos-threat-report-2026-h1/

Najafimehr, M., Zarifzadeh, S., & Mostafavi, S. (2023). DDoS attacks and machine-learning-based detection methods: A survey and taxonomy. Engineering Reports, 5(12), e12697. https://doi.org/10.1002/eng2.12697

Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., & Hamin, M. (2025). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations (NIST AI 100-2e2025). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-2e2025