Countermeasures in IoT Devices Using Reinforcement Learning

César Rodríguez Villagrá[1], Amanda Pérez Olmos[1], Rubén Ruiz
González[1], and Nuño Basurto[1]

Grupo de Inteligencia Computacional Aplicada (GICAP), Universidad de Burgos,
Av. Cantabria s/n, 09006 Burgos, Spain
{crv1002, apo1004}@alu.ubu.es, {ruben.ruiz, nbasurto}@ubu.es

Abstract. Cybersecurity in IoT devices is crucial for ensuring continuous functionality and service availability. As these devices become more interconnected, they are becoming increasingly vulnerable to cyberattacks, particularly Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) attacks, which have caused severe disruptions in recent years. These attacks can overload the network infrastructure, causing slow performance or total service failures. To address this issue, the implementation of a reinforcement learning (RL) agent within the network router is proposed to intelligently detect and mitigate DoS and DDoS attacks in real-time. Unlike traditional rule-based security mechanisms, RL agents dynamically adapt to evolving attack patterns, making them more effective against sophisticated threats. The agent learns to identify malicious traffic by attempting to saturate the queue and discard it, preventing network congestion and ensuring stable performance. The key objective of this approach is to minimize packet loss while maintaining high network efficiency. The experimental results demonstrate that the proposed system effectively reduces the impact of DoS and DDoS attacks, strengthening the resilience of IoT networks against cyber threats.

Keywords: Internet of Things · Security · Reinforcement learning.

1 Introduction

The rapid growth of the Internet of Things (IoT) has introduced new challenges for ensuring the security of networked devices. As IoT devices often have lower security measures, they are prime targets for cyber-attacks, including Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks [2] [4], which disrupt the functionality of devices within the network. It is crucial to ensure security through the application of advanced techniques, such as Reinforcement Learning (RL), a branch of artificial intelligence (AI) that offers a promising approach to dynamically detect and mitigate these threats [1] [3].

2 Experimental Setup

Python was used for synthetic data generation, creation of the environment, training of the agent, and visualization of progress and statistics. The main li-

braries are Gymnasium, which utilizes existing reinforcement learning knowledge to create a customized model, and Stable-Baselines3 for training, loading, and saving agents. The simulation environment was designed to model a simplified router that attempted to mitigate attacks with the following features.

- **Observation Space**: Represent the occupation of the queue (in %) and the number of discarded packets.
- **Action Space**: Only packets can be permitted or denied.
- **Reward Function**: Focuses on the discarded packets and the actual action. Trying to minimize discarded packets.

## 3   Reinforcement Learning for DoS and DDoS mitigation

The mitigation of cyber-attacks in traditional networks is a challenging task and, at the same time, essential for ensuring the security and availability of services in modern digital infrastructures. In the initial phase of this work, the focus was on mitigating one of the most common threats, DoS and DDoS attacks. In this section, the initial steps for designing the environment, training, and deploying an RL agent to mitigate these attacks are explained.

3.1   Environment

The experiment needs:

- **Packet Source Generation**: Currently, a simulated packet generation, but it could be from the internet or another local area network (LAN).
- **Router**: In this case, the simulated router has a single queue for packets.
- **Agent**: The RL agent decides whether to drop or allow the incoming traffic at each time step.
- **IoT LAN**: IoT devices that receive the packets.

Figure 1 illustrates the theoretical network architecture used in the simulations.
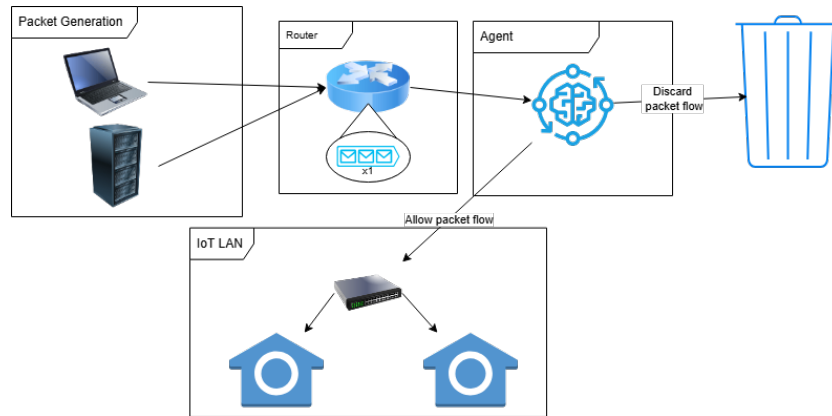
Fig. 1. Theoretical Environment set-up.

3.2 Agent Training

For training, the PPO Stable-Baselines3 implementation was used with a Multi-Input Policy, 200,000 time steps, and default parameters. Agents learn with only the information of % of the queue occupation and discarded packets in the actual step, and have to choose between allowing or denying all traffic in the step.

4 Experimental Results

At this stage, the agent is trained with only DoS attacks and benign traffic. As shown in Figure 2 the agent learns in response to the mean reward.

4.1 Training Metrics

Fig. 2. Reward and explained variance along training.

The reward mean is a key metric in RL, providing insight into an agent's performance during training. A consistent increase in the reward mean typically indicates that the agent is improving its decision making policy. Another important metric in RL is explained variance, which measures how well the model value predictions correlate with the actual rewards received. It helps identify whether the value function is learning effectively or requires further optimization, offering a quantitative measure of the model's predictive power.

4.2 Agent Metrics

Figure 3 shows the behavior of the system during an attack. As the queue approaches its capacity due to the incoming attack traffic, the agent's action changes. It begins to deny packets, effectively reducing the load on the network

until the attack intensity decreases and the queue stabilizes. This dynamic response illustrates how the agent adapts its decisions to mitigate the impact of DoS or DDoS attacks.
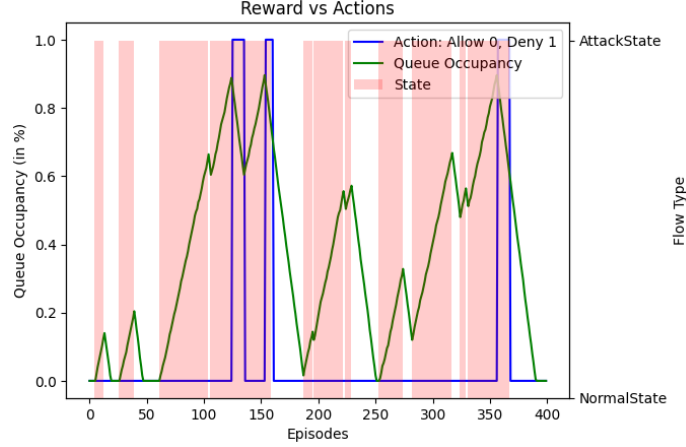
Fig. 3. The queue occupation and the action in each step.

5   Conclusions and Future Work

The increasing use of IoT devices has made them more vulnerable to cyberattacks, particularly DoS and DDoS attacks, which can severely disrupt network performance. This study demonstrates that reinforcement learning can effectively mitigate these threats by dynamically adapting to evolving attack patterns. The proposed RL-based system reduces the impact of these attacks on the local network, enhancing its stability and security.

Future work will focus on improving the adaptability of the RL agent to more sophisticated and diverse attack strategies, such as PortScan and OSScan. In addition, the expansion of the system's capabilities to detect and mitigate other network threats will be explored.

References

1. Gasmi, R., Hammoudi, S., Lamri, M., Harous, S.: Recent reinforcement learning and blockchain based security solutions for internet of things: Survey. Wireless Personal Communications **132**, 1307–1345 (9 2023). https://doi.org/10.1007/s11277-023-10664-1

2. Liu, Y., Dong, M., Ota, K., Li, J., Wu, J.: Deep reinforcement learning based smart mitigation of ddos flooding in software-defined networks. In: IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, CAMAD. vol. 2018-September (2018). `https://doi.org/10.1109/CAMAD.2018.8514971`
3. Sewak, M., Sahay, S.K., Rathore, H.: Deep reinforcement learning in the advanced cybersecurity threat detection and protection. Information Systems Frontiers **25** (2023). `https://doi.org/10.1007/s10796-022-10333-x`
4. Yungaicela-Naula, N.M., Vargas-Rosales, C., Pérez-Díaz, J.A., Carrera, D.F.: A flexible sdn-based framework for slow-rate ddos attack mitigation by using deep reinforcement learning. Journal of Network and Computer Applications **205** (2022). `https://doi.org/10.1016/j.jnca.2022.103444`