Cybersecurity Intrusion Response Systems using Reinforcement Learning

Rogelio Velázquez
Associate Professorship of Embedded Systems and Internet of Things
Technical University of Munich
Munich, Germany
ge57piw@mytum.de

***Abstract*- The variety of cyber-security incidents in our modern society around the world is constantly evolving due to the fast pace of technological development. Vulnerabilities in various cyber-physical systems can compromise critical information in a matter of seconds. Therefore, the development of systems that can react to potential attacks must be at the forefront to identify malicious intentions and mitigate the effects that they could have if these can break through the security barrier of our system. The main objective of this paper aims to introduce Intrusion Response Systems (IRS) that employ Reinforcement-Learning (RL) to select a desired reaction for cyber-attacks in IRS. An evaluation of the proposed IRS is provided considering their general architecture and techniques for decision making. The improvements of using RL in IRS above other approaches are shown and the key distinctions of IRS among other related concepts are also stated.**

***Keywords*—** *Reinforcement Learning, Intrusion Response Systems, Deep Learning, Mitigation, Response, Denial of Service (DoS), Distributed Denial of Service (DDoS).*

I. Introduction

Whenever any device connects to the internet, it becomes susceptible multiple threats and risks on the web. The Internet is a vast network that connects millions of devices all over the world. The volume of data traffic on the web is immense, and it will keep growing. Unfortunately, it is not possible to control the malicious intentions of certain persons, which creates opportunities to exploit valuable information that is found online.

According to the IV-Test Institute of IT-Security of Germany, only in 2022 around the world (for Windows Users), 791,882,554 Malware were documented along with 188,227,795 Potentially Unwanted Applications. More information and statistics for MacOS, Linux, and Android users is shown on [1]. As regards the EU, the CIRAS incident reporting [5] shows that in 2022, critical service providers notified of 1083 incidents with a significant impact on the national authorities of their respective countries. This means that valuable assets like servers, domain controllers, websites, etc. were affected by different kinds of cyber-attacks. It is worth mentioning that some types of attacks are getting more common, the European Union Agency for Cybersecurity states in its Threat Landscape report of 2022 [2] that DDoS attacks are increasing and getting more complex, targeting mobile networks and Internet of Things (IoT) devices, playing a role in cyber warfare.

This paper highlights a tight relationship between IRS and DDoS attacks. It is not a coincidence that many of the available sources that can be found relating to this topic usually propose models that react against this type of attack. For example, [4] a study conducted by Bitkom e.V., the industry association for the German information and telecommunications sector surveyed around 1,000 companies and revealed that 27% experienced damage from DDoS attacks between 2020 and 2021. This ranked DDoS attacks as the second most damaging type, behind malware at 31%.

This paper provides and introductory review of IRS that use RL to mitigate cyber-attacks. The main contributions of this work are: (1) compare different IRS approaches that use RL to react against a given attack, by highlighting the main components of RL for selecting a response, like the learning algorithms and the architecture used for the proposed models, as well as going through the main considerations for the use of RL, such as attack scenarios or the training of the model. (2) provide demonstrations of how RL approaches can outperform the mitigation efficiency above other techniques. The purpose of these contributions is to show that RL can have efficient reactions against certain types of attacks, especially DDoS.

The paper is structured as follows: section II. covers basic foundations, section III. discusses related work, section IV. describes the proposed method, section V. presents the evaluation of selected models, and the conclusions are in the final section.

II. Background

The reviewed proposals base their approaches on general concepts that are explained below:

1) Intrusion Response Systems: IRS can be defined as a security countermeasure that is triggered when a system intrusion is detected. The main goal of IRS is to take actions in case that an attack is detected to mitigate its effect on the protected system. The design of this systems is done by considering different parameters, according to the desired goals of each application. These systems must be able to collect and process information to provide alerts over detected attacks and to provide mitigation responses to those attacs. IRS can be categorized regarding their design parameters according to [12], [13] and, [14]. Depending on their automation level they can be divided into notification, manual, and automatic IRS. All of the proposals reviewed in this paper are automatic IRS.

2) Reinforcement Learning: The meaning of "Learning" refers to the process of acquiring new knowledge by a certain mean (experience, studyin, etc.)[6]. Machine Learning (ML) learns by a process called inductive reasoning to acquire knowledge through inference. ML relies on probabilistic claims and observed patterns to make predictions about future situations. RL "learns" by interacting with the environment in which is immersed [7]. The main components of an RL model are seen below:

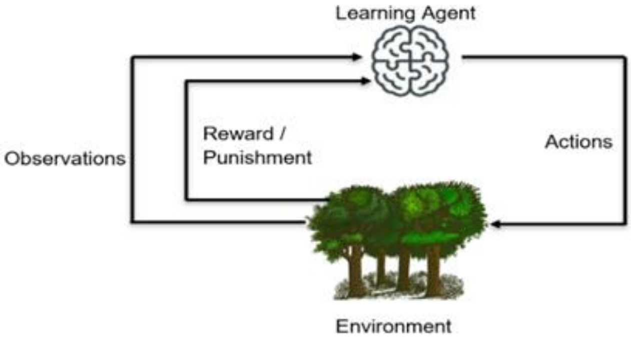

Figure 1. The basic components of an RL model

Where each component has the following functions:

- Learning agent: Entity that interacts with the environment and takes actions based on states and receives feedback in the form of rewards or penalties.
- Environment: External system with which the agent interacts to learn, providing states to the agent, and giving feedback based on the agent's actions.
- Reward/punishment: Feedback received by the agent from the environment after an action is taken.
- Actions: Decisions taken in response to a given state.
- Action space: refers to the set of all possible actions in an environment).
- In an RL model we will usually also have:
- State: Relevant information for the agent's decision-making.
- Policy: The strategy followed to map states to actions and guide the agent's decision-making process. It can be deterministic or stochastic.
- Value function: Estimation of the expected cumulative rewards from a state or state-action pair.
- State space: Set of all possible states that an RL agent can encounter in the environment.

An RL model will usually seek the agent's goal to be to maximize the cumulative rewards.

3) Deep Learning and Deep Reinforcement Learning (DRL): Deep learning is an algorithmic approach that uses neural networks modeled after the human brain. (AWS, 2023). Deep Learning Neural Networks (DNN) consist of artificial neurons that process data through mathematical calculations. The main components of a deep learning network are the next ones:

- Input layer: Data that goes into the network.
- Hidden layer: Processes and transfers the data to the next layers on the network.
- Output layer: Nodes that output the data.

DRL is basically a combination of RL and Deep Learning, using DNN as function approximators. Deep Q-Networks (DQN) are DNNs that utilize the Q-learning algorithm. The Q algorithm estimates expected rewards for actions in each state.

III. Related Work

The state of the art regarding IRS is not particularly as broad as other related areas, like Intrusion Detection Systems, which focus on the detection process of cyber threats. Nevertheless, research interest has also been shown in different approaches that use RL to trigger mitigation effects after certain cyber attacks are detected on the system.

In this review, it was found that DDoS is a common type of attack in which related works mainly focus. This paper reviews 4 different proposals that use RL approaches to select an appropriate response against a cyber attack. These were selected due to their similarities in their architecture components and their focus on defending against similar types of attacks. [9] Proposes a framework for defending against various DDoS attacks, [10] presents an IRS for DoS attacks, and [11] targets defense against DDoS attacks as well. For extending to other kinds of threats, a different approach is also considered in [11] where they propose a game theory-based algorithm to extend RL in IRS to other threat types. Despite using different RL approaches, a correlation of their architectures could be done, and a better performance against other approaches is also demonstrated in these papers. This paper also reviews the use of different algorithms in the proposed IRS. This job also focuses on correlating the key aspects between the models taken as reference, providing a brief explanation for each of them. The performance improvement for IRS using RL above other techniques is also shown. The proposals evaluation is presented in more detail in the next section.

IV. Evaluation

This study aims to extract key information from a reduced number of proposals, connecting related concepts and presenting them in a concise way. This paper provides an introductory understanding of IRS that use RL against cyber-attacks, unlike other papers that delve into exhaustive surveys of numerous proposals and require an advanced technical knowledge and experience in the field.

After performing a comprehensive analysis of key aspects of reference proposals, certain general features have been identified. These features are summarized in the next points and addressed afterward with more detail:

1) *Model inputs:* Inputs for the RL model to collect information. These come from different information sources, which can include detection tools such as Intrusion Detection Systems (IDS). The collected information assists the IRS in selecting a desired response.
2) *Basic Architecture:* General architecture of an IRS that uses RL for selecting a response after an attack. The architectures of the reviewed models link the information collection module and the mitigation module to have an effect on the targeted system.
3) *Algorithms:* Certain algorithms that work well with RL models are used for decision making processes and response optimization of the IRS.
4) *Responses:* Possible responses that an IRS can have. These can be either passive or active.
5) *Performance:* Performance (in terms of percentage of effective responses) of proposals against other

approaches and general considerations for RL models.

It is important to highlight that although these approaches target to defend against a similar attack, their design strongly differs from each other. Therefore, it is not possible to weigh each aspect of the proposals (such as the algorithms used or their responses) between each other, as they meet different needs. The relation of the similarities between them is exposed for comprehensive purposes.

1) *Model inputs:* RL models need inputs for the decision-making process. Traffic information needs to be collected so that the RL agent can use it to take proper actions. In [9] they collect data using the standard protocol OpenFlow and they define the central agent's state space: It consists of 8 features derived from statistics out of an SDN protocol (OpenFlow used as the SDN protocol and source for the features) such as the port number, packet count, etc. The state space is directly related to the action space. The other proposals [7], [8], [9], collect traffic information as well, but they do not use a standard protocol for this.

2) *Basic Architecture:* A correlation between the general components integrated into the architectures of the reviewed proposals was done. In the next image, an architecture of the analyzed IRS that integrates the main components of an RL is proposed. This was correlating the architectures proposed at [6], [7], [8], and [9]:

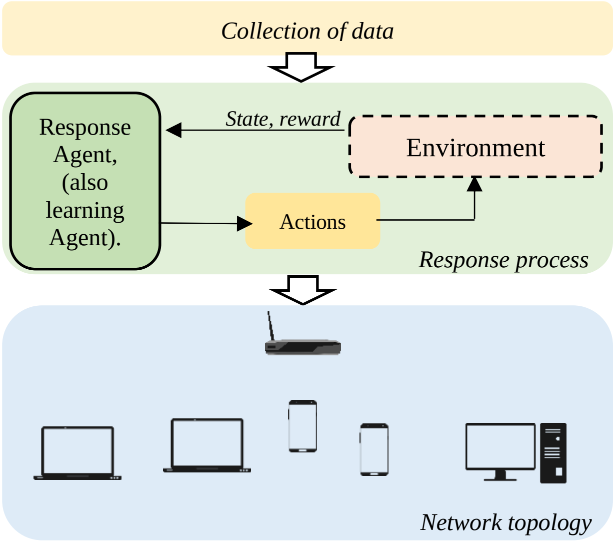

Figure 2. General architecture of an IRS that uses RL to respond to an attack.

The previous image illustrates the main components that correspond to a general architecture for the IRS proposed by the authors. 3 main processes are identified, which is the collection of data, the response process, and the network topology. In the collection of data, information is gathered to feed the model with inputs as network traffic or datasets for training the model. In the response process, the RL model gets the information from the environment (data collected from process 1.), sets a state for the Response Agent, and then the Agent responds with an action depending on the state that was set. The action needs to have effect on the environment and if the action matches the expected goal stablished for the model, then the Agent receives a reward. If the action does not match with the expected goal, the agent will ideally receive a "punishment" as a corrective measure to an inappropriate response. The response to the attack is provided by an entity. In each proposal they have response agents that fulfil different their needs according to their design. In [9] they integrated a learning agent into a "defender router". In [10] they call it "control agent". In [11] they call it "mitigation engine".

3) *Algorithms:* RL models require decision-making algorithms to determine optimal which are the most optimal actions to take regarding different states of the environment. These algorithms serve to optimize the agent's actions based on learned knowledge to improve the decision-making process over time.

In [9] they proposed an actor-critic-based DDPG (Deep Deterministic Policy Gradient) algorithm, which defines the current policy and maps states to specific actions deterministically. In [10] they use a Q learning algorithm to obtain the optimal policy and a Double Q-learning algorithm for quickly obtaining the optimal policy while defending against attack traffic flows (reducing the impact on benign traffic as well). In [11] they use a discrete-time partially observable Markov decision process (POMDP) based on game theory, involving 2 players, the attacker, and the defender.

In each method, the algorithm's policy purpose is to describe how an agent behaves. This maps the states to probabilities of choosing an action. "Obtaining the policy" refers to the process of learning and generating an optimal policy to achieve a given goal. The process of obtaining the policy requires repeated interactions with the environment, updating the policy based on the obtained rewards and improving its decision-making accuracy.

4)*Responses:* The most important part of the IRS is the mitigation that these provide. Active responses change the system's state to block a detected attack and mitigate the effects that this could have. On the other hand, passive responses only raise alerts and notify operators. [12]. The action space (related to the state space) is notoriously tied to deciding what to do after an attack. For example, in [9] each defender router applies throttling of the traffic through OpenFlow meters. They defined that if the aggregate traffic rate exceeds the maximum bandwidth, excess traffic will be dropped, and that benign traffic should not be dropped to not let attackers reach their goals. In [10] they deploy an optimal response policy to tackle malicious traffic flows while minimizing the impact on the forwarding of legitimate traffic flows. A different response is proposed in [11], where they use a game theory, defining an attacker, and a defender. Here the defender keeps a belief about the network state and the attacker's strategy based on information from the attack graph, network event alerts and previous mitigation actions. The defender calculates probabilities for each security condition being enabled and matches network alerts to specific nodes in the attack graph. Based on the belief about the network state, the defender responds to an attack with defensive actions, like deploying IPtables firewall rules to block exploit nodes in the attack graph.

5) *Performance:* All the proposals results are obtained in different environment setups, using different tools and different algorithms. Each paper provides a comparison of the methodology developed against certain state-of-the-art techniques that are different for each paper. However, all IRS-

RL-based approaches outperform the techniques to which they are compared.

In [9] improvements of the responses are shown by forwarding a higher percentage of benign traffic and blocking more of the malicious traffic. The authors compare this IRS-DRL based approach against 2 other approaches: a popular router throttling technique called Additive Increase Multiplicative Increase (AIMD) and a Coordinated Team Learning approach. 5 different attack patterns were set, and the mitigation system didn't learn these patterns. Each attack pattern is conducted for 100 episodes. The DRL outperforms the other two approaches in both benign traffic forwarding and in attack. In [10], the proposed model shows the defense performance by measuring the ratio of dropped malicious packets since an attack is detected. 3 DoS attack scenarios are set. "DeepAir" which is the name of the developed approach, prevents 85% of the attack packets from arriving to the victim. The Q-learning approach prevents 60% of the attack packets arriving. GATE and GTAC-IRS prevent a significantly lower number of packets, only 30% and 45%, respectively. In [11], they consider sixteen different configurations of the IRS to evaluate the mitigation capabilities. They tested thirteen different cyber-attacks scenarios, setting a Firewall rule type (either Specific or global) and they measured whether the attack was mitigated or not. It is noteworthy that in this approach they integrate an Intrusion Detection System (IDS) to alert the IRS.

It is important to highlight that although words [9], [10], and [11] share similar characteristics, their modeling is significantly different from each other. The purpose of this review is not to compare these proposals against each other, because they target their system's protection in different ways, and the responses they provide are also different. On the other hand, this review remarks the strong relationship between the structures of different IRS that use RL to select a response. This review also highlights that these approaches can outperform other state-of-the-art techniques that react with an equivalent protection. It is also important to consider that each model is trained differently from each other, and that the inputs for training the model will usually be dependent on the type of attack and the desired response.

V. Conclusions

This work links the RL algorithms in IRS used to improve the efficiency of mitigation responses against given attack(s). of the model for certain situations among other things. Deeper considerations need to be taken like the optimization of the model for certain situations among other things. This will be strongly related to the type of responses that the model needs to have, the environment with whom it interacts, the network assumptions taken for the model, how is trained, and other factors to contemplate.

It is worth mentioning that although the review of these proposals shows the benefits of the mitigation effects of the IRS responses, when modelling an IRS to select a response against an attack it should not be considered only the design of the model (i.e., the model's architecture, the RL algorithm, etc.) but also external factors such as environment limitations and how the IRS can be integrated into the system. Nevertheless, vulnerabilities in these proposals were also detected, such as low or zero information about unknown attacks, the distribution of data in real world-scenarios could differ a lot from the training data, exploitation of RL algorithms, and even the use of ML to have more efficient attacks (like adversarial learning attacks, which also use Machine-Learning based attacks).

References

[1] AV-Test, the Independent IT-Security Institute. (2023). Total Reported Malware in the world graph. Retrieved from https://www.av-atlas.org/

[2] Federal Office for Information Security. (2022). The State of IT Security in Germany in 2022. Retrieved from https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Securitysituation/IT-Security-Situation-in-Germany-2022.html?nn=1021082

[3] European Union Agency for Cybersecurity (ENISA). (2022). ENISA Threat Landscape (ETL). Retrieved from https://www.enisa.europa.eu/topics/cyber-threats/threats-and-trends

[4] Bitkom e. V. (2021). Angriffsziel deutsche Wirtschaft: Mehr als 220 Milliarden Euro Schaden pro Jahr [Attacking the German Economy: More than 220 Billion Euros in Damage per Year]. Retrieved from https://www.bitkom.org/Presse/Presseinformation/Angriffsziel-deutsche-Wirtschaft-mehr-als-220-Milliarden-Euro-Schaden-pro-Jahr

[5] European Union Agency for Cybersecurity (ENISA). (2022). Ciras Incident Reporting. Retrieved from https://ciras.enisa.europa.eu/

[6] Dávila Newman, G., (2006). El razonamiento inductivo y deductivo dentro del proceso investigativo en ciencias experimentales y sociales. [Inductive and Deductive Reasoning in the Investigative Process in Experimental and Social Sciences]. Laurus, 12(Ext),184-188.

[7] Fraunhofer IKS. (2021). Munich The components of the Reinforcement Learning framework. Retrieved from https://safe-intelligence.fraunhofer.de/en/articles/a-quick-look-at-safe-reinforcement-learning

[8] H. Alturkistani and M. A. El-Affendi, "Optimizing cybersecurity incident response decisions using deep reinforcement learning" in *International Journal of Power Electronics and Drive Systems (IJECE).* IJECE, December 2022. [Online]. Available: https://doi.org/10.11591/ijece.v12i6.pp6768-6776

[9] Y. Liu, M. Dong, X. Jia, J. Li and Wu, J. (2018). "Deep Reinforcement Learning based Smart Mitigation of DDoS Flooding in Software-Defined Networks" in *IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)* IEEE, 2018. [Online]. Available: https://doi.org/10.1109/camad.2018.8514971

[10] T. Q. Phan and T. Bauschert. (2022). "DeepAir: Deep Reinforcement Learning for Adaptive Intrusion Response in Software-Defined Networks" in *IEEE Transactions on Network and Service Management Transactions on Network and Service Management, vol. 19 no. 13.* IEEE. September. 2022. [Online]. Available: https://doi.org/10.1109/tnsm.2022.3158468

[11] J. R. Rose, M. Swann, K. P. Grammatikakis, I. Koufos, G. Bendiab, S. Shiaeles and N. Kolokotronis, "IDERES: Intrusion detection and response system using machine learning and attack graphs" in *Journal of Systems Architecture*. Elsevier Ltd. 2022. [Online]. Available: https://doi.org/10.1016/j.sysarc.2022.102722

[12] May Bashendy, Ashraf Tantawy, Abdelkarim Erradi, Intrusion Response Systems for Cyber-Physical Systems: A Comprehensive Survey, Computers & Security (2022) DOI: https://doi.org/10.1016/j.cose.2022.102984

[13] Z. Inayat, A. Gani, N. B. Anuar, M. K. Khan, and S. Anwar, "Intrusion response systems: Foundations, design, and challenges," in *Journal of Network and Computer Applications*. Elsevier Ltd. 2016, pp. 58-67. [Online]. Available: https://dx.doi.org/10.1016/j.jnca.2015.12.006

[14] S. Anwar, J. M. Zain, M. F. Zolkipli, Z. Inayat, S. Khan, B. Anthony, and V. Chang, "From intrusion detection to an intrusion response system: Fundamentals, requirements, and future directions," in *MDPI Journal.* MDPI , March 2017, pp. 15-18. [Online]. DOI: 10.3390/a10020039

[15] A. Shameli-Sendi, M. Cheriet, and A. Hamou-lhadj, "Taxonomy of intrusion risk assessment and response system," in *Computers and security*, vol. 5, pp. 8–11, 2014. [Online]. Available: https://doi.org/10.1016/j.cose.2014.04.009