AI and Data Science Journal
ISSN: 3078-7394
DOI: https://doi.org/10.61784/adsj3010

AUTOMATED CYBERSECURITY INCIDENT RESPONSE: A REINFORCEMENT LEARNING APPROACH

MingJie Zhao, Rui Chen*
Beijing University of Posts and Telecommunications, Beijing 100000, China.
Corresponding Author: Rui Chen, Email: rchen272@bupt.edu.cn

Abstract: Cybersecurity incident response is critical for defending digital infrastructures from evolving cyber threats. Traditional manual systems and rule-based automation methods cannot efficiently cope with dynamic and sophisticated attacks. This paper explores the application of reinforcement learning (RL) to automate cybersecurity incident response. By modeling the response process as an RL problem, where an agent learns from interactions with the environment, the proposed system aims to enhance detection accuracy, minimize response times, and reduce false positives. Experimental results demonstrate the system's ability to mitigate threats effectively, showing that RL can significantly improve the efficiency and scalability of cybersecurity defenses. This approach leverages machine learning to automate decisions in real-time, adapting to evolving threats and optimizing incident response strategies. The integration of RL in incident response has the potential to dramatically reduce human error, improve system adaptability, and scale efficiently in complex, high-volume environments.
Keywords: Cybersecurity; Incident response; Reinforcement learning; Automation; Cyber threats; Machine learning; AI

1 INTRODUCTION

In the contemporary digital era, cybersecurity has become an urgent priority for individuals, businesses, and governments. As organizations increasingly depend on digital platforms to store sensitive information and conduct transactions, cyber threats have evolved to become more sophisticated, diverse, and dangerous. Ransomware, phishing attacks, and Distributed Denial-of-Service (DDoS) attacks are just a few examples of the growing cyberattack landscape that affects various sectors globally. These attacks, if left unchecked, can lead to severe financial loss, data breaches, reputational damage, and even national security threats[1].

Traditional cybersecurity methods, such as signature-based detection systems, firewalls, and antivirus software, have served as the foundation of digital security for decades. However, these methods are increasingly failing to keep up with the dynamic and evolving nature of modern cyberattacks. Signature-based systems, for instance, are effective only against known threats and are often inadequate for detecting new, unseen, or polymorphic attacks. Rule-based systems, though effective for predefined scenarios, lack the adaptability to address novel threats in real time[2]. Moreover, the increasing complexity of attacks requires more than just automated responses; it demands intelligent, adaptive decision-making systems that can optimize their response strategies as new threats emerge.

Incident response, which is a critical component of cybersecurity, involves identifying, analyzing, and mitigating security breaches as quickly and effectively as possible. Traditional incident response often depends on manual intervention, human judgment, and predefined response strategies. However, as cyber threats grow in volume and complexity, human intervention becomes slower, more error-prone, and increasingly inadequate to keep up with the pace of modern attacks. Furthermore, rule-based systems and predefined responses can be rigid and fail to detect or properly respond to new types of attacks. Thus, there is a growing need for automated, scalable, and real-time incident response systems that can handle the growing number of cybersecurity incidents without compromising performance.

RL, a subfield of machine learning, provides a promising solution to these challenges. Unlike traditional rule-based systems, RL enables systems to learn from experience, make decisions autonomously, and adapt to new threats over time. RL models are based on a feedback loop, where an agent interacts with an environment, takes actions, and receives feedback in the form of rewards or penalties. Over time, the agent learns to make optimal decisions that maximize cumulative rewards. In the context of cybersecurity incident response, this means that the RL agent can continually learn from past interactions, improving its decision-making and adapting to new types of cyber threats [3].

This paper explores the application of RL to automate cybersecurity incident response. We propose a system architecture that leverages RL for real-time threat detection and mitigation. Our system continuously learns from feedback and adapts to evolving attack patterns, offering a more dynamic and effective solution than traditional methods[4]. Through experiments and simulations, we demonstrate that RL-based incident response systems can outperform rule-based systems in key areas such as detection accuracy, response time, and false positive rates. This paper also discusses the advantages of RL in providing a scalable, adaptive, and efficient solution for modern cybersecurity challenges.

2 LITERATURE REVIEW

Cybersecurity is a critical domain that requires constant innovation due to the ever-evolving nature of cyber threats. Traditional approaches to cybersecurity largely rely on manual intervention or rule-based systems, which focus on predefined attack signatures or patterns. These systems work well for known threats but struggle when faced with new, previously unseen attack types[5]. Over the years, several machine learning techniques have been explored to address the limitations of traditional methods, and they have shown great promise in enhancing the efficiency and effectiveness of cybersecurity systems.

In particular, machine learning (ML) has become an essential tool in the cybersecurity field due to its ability to detect new, unknown attacks that may not fit predefined patterns. Various ML techniques, such as supervised learning, unsupervised learning, and deep learning, have been applied in areas like intrusion detection, malware analysis, and anomaly detection. Supervised learning, for example, requires labeled data for training, where the system is taught to identify specific features of known threats[6]. It has been successfully applied to malware classification and phishing detection, where the model learns from labeled examples of malicious and benign behavior.

However, supervised learning models face limitations, particularly when it comes to detecting unknown threats. Since these models rely on previously labeled examples, they are incapable of identifying new, unseen threats unless retrained with new labeled data. This dependency on labeled data makes supervised learning less practical in dynamic environments like cybersecurity, where new threats emerge continuously and require the system to adapt in real time. Unsupervised learning, on the other hand, does not require labeled data and can detect anomalies or deviations from normal behavior, making it useful for identifying novel attacks. However, unsupervised learning models often result in a higher rate of false positives, which can undermine the overall effectiveness of the system.

Deep learning, which is a subset of machine learning that uses artificial neural networks to learn from large amounts of data, has gained attention in the cybersecurity domain. Deep learning models have been applied to tasks such as intrusion detection and malware classification[7]. These models are capable of identifying complex patterns in large datasets and are particularly effective at handling unstructured data, such as network traffic or raw log files. However, deep learning models require substantial amounts of training data and computational power, which can be a limiting factor in resource-constrained environments.

RL, a form of machine learning in which an agent learns to make decisions by interacting with its environment, has emerged as a powerful tool in dynamic and sequential decision-making tasks. Unlike supervised and unsupervised learning, RL does not rely on labeled data or predefined patterns. Instead, the agent learns by taking actions and receiving feedback in the form of rewards or penalties[8]. This makes RL particularly well-suited for environments like cybersecurity, where attack patterns are constantly evolving, and predefined rules or signatures are insufficient.

RL has been applied to various cybersecurity tasks, such as intrusion detection, network defense, and malware analysis. In intrusion detection systems (IDS), RL agents have been trained to classify network traffic as benign or malicious based on patterns learned from past interactions. Similarly, RL has been used to optimize strategies for network defense, where the agent learns to block malicious traffic, isolate infected systems, or adjust security configurations in response to ongoing attacks. In malware analysis, RL has been applied to identify malicious behavior by observing how software interacts with the system and learning to distinguish between benign and malicious activities.

Despite the promising results in these applications, the full-scale application of RL to automate complete incident response systems—encompassing detection, decision-making, and mitigation—remains an area of active research. While RL offers the potential to significantly improve the efficiency and adaptability of cybersecurity systems, there are several challenges that must be addressed before RL can be widely implemented in real-world cybersecurity environments. One of the primary challenges is the need for high-quality training data. RL agents rely on feedback from interactions with the environment to learn optimal decision-making strategies. If the environment is not accurately modeled or the data is biased, the RL agent may learn suboptimal or even harmful strategies.

Another challenge is the computational complexity of RL models. Training RL agents requires significant computational resources, particularly when working with large-scale datasets or complex environments. For instance, training a deep RL agent to handle real-time cybersecurity incidents requires massive computational power, which can be a limiting factor in environments where resources are constrained. Additionally, the time it takes to train an RL model can be a barrier, as cyberattacks evolve rapidly and require systems to adapt almost instantaneously.

Furthermore, RL-based systems are susceptible to adversarial attacks. Adversarial attacks aim to manipulate the learning process of RL agents by introducing misleading or malicious feedback[9]. These attacks can potentially compromise the security of the system, allowing attackers to bypass detection or manipulate the system's decision-making process. To address this vulnerability, it is essential to develop techniques to secure RL systems against adversarial manipulation and ensure their robustness in the face of such threats.

Despite these challenges, the application of RL in cybersecurity incident response has the potential to revolutionize the field by enabling systems to continuously learn, adapt, and improve over time. By automating decision-making and optimizing responses to emerging threats, RL-based systems can reduce response times, improve detection accuracy, and scale more effectively to handle large volumes of cybersecurity events. As RL research continues to advance, it is expected that RL will play an increasingly important role in the future of automated cybersecurity.

3 METHODOLOGY

3.1 System Architecture

The proposed system architecture for the RL-based automated incident response is designed to address the dynamic and evolving nature of cybersecurity threats. Traditional rule-based systems, while effective against known threats, lack the flexibility and adaptability to handle novel or complex attacks. This is where RL offers significant advantages by providing the system with the ability to learn from experience and adapt its responses to new and unforeseen threats. The architecture of the RL-based incident response system is composed of several key components, which are crucial for the efficient and accurate identification, analysis, and mitigation of cybersecurity incidents.

The first component of the system is data collection, which is critical for providing the RL agent with the information it needs to make informed decisions. This data is gathered from multiple sources, such as network traffic, system logs, and external threat intelligence feeds. Network traffic data includes information on packet flow, source/destination IP addresses, port numbers, and timestamps, which can provide crucial indicators of potential cyber threats. System logs capture detailed information about system events, user actions, and application behavior, which can also help identify suspicious activities that may indicate an ongoing attack. External threat intelligence feeds provide valuable information on known attack patterns, emerging threats, and vulnerabilities, which can further assist in identifying potential risks.

Once the data is collected, it is preprocessed and analyzed to extract relevant features that can provide insights into the state of the system and potential threats. Preprocessing involves filtering out noise from the data, normalizing the data, and identifying patterns or anomalies that are indicative of malicious activities. This is often achieved using machine learning algorithms such as clustering or anomaly detection techniques, which are designed to recognize deviations from normal behavior. For example, if a particular IP address is sending an unusually high volume of requests to a server, the system may flag this as a potential denial-of-service (DoS) attack. Similarly, if there is suspicious behavior or unauthorized access attempts detected in system logs, it may indicate a potential breach.

After the data has been processed and potential threats have been identified, the next step involves the RL agent, which is the heart of the automated incident response system. The RL agent is responsible for analyzing the current state of the system and deciding on the most appropriate action to mitigate the detected threat. The decision-making process is based on an evaluation of the current system status, which includes factors such as the nature and severity of the threat, the potential impact on the system, and the available resources to respond to the attack. The RL agent learns to take optimal actions through interactions with the environment, where it continuously improves its decision-making based on feedback received from past actions.

In an RL framework, the agent makes decisions through a process known as policy learning. The agent's policy is a strategy that determines the best action to take in each state. The policy is refined over time as the agent interacts with the environment and receives feedback. The feedback comes in the form of rewards and penalties, which help the agent learn which actions are effective in mitigating threats and which are not. For example, if the RL agent successfully blocks a malicious attack, it receives a positive reward, whereas if it fails to detect or mitigate the attack in time, it receives a penalty. This process of trial and error allows the agent to learn from its mistakes and continuously improve its policy.

Once the RL agent determines the optimal action to take, the system moves to the final stage, which is response orchestration. This component is responsible for executing the mitigation actions decided by the RL agent in real-time. The system can carry out various actions, such as blocking malicious network traffic, isolating infected systems from the network, or alerting administrators for manual intervention if necessary. For example, if the RL agent detects a ransomware attack, the system might isolate the infected host and block any outgoing traffic to prevent further spread of the malware. In the case of a DDoS attack, the system may reroute traffic or block malicious IP addresses to minimize the impact of the attack.

The entire architecture is designed to operate in real-time, ensuring that the system can respond to cybersecurity threats as quickly as possible. The RL agent is continuously learning from new data and adapting to emerging threats, allowing the system to handle a wide range of attack scenarios. Moreover, the system is highly scalable, enabling it to manage large volumes of security events without compromising performance as in Figure 1.

Figure 1 System Architecture of RL-Based Automated Incident Response

3.2 Reinforcement Learning Model Design

The RL model used in the system is based on the Markov Decision Process (MDP) framework, which provides a mathematical model for decision-making under uncertainty. In this framework, the state represents the current conditions of the system, such as active threats and network activity. The actions correspond to the possible responses that the RL agent can take, such as blocking traffic or isolating a system. The reward function provides feedback on the effectiveness of the actions taken. Positive rewards are given for successful mitigation, while penalties are assigned for failures.

To train the RL model, we use Q-learning, a model-free reinforcement learning algorithm. Q-learning allows the agent to update its decision-making policy based on the rewards or penalties it receives from the environment. The agent learns to maximize cumulative rewards by selecting optimal actions in each state.

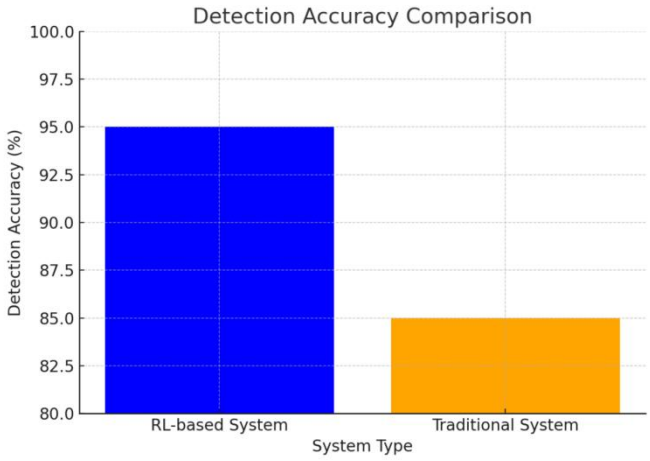
Q-learning involves updating Q-values, which represent the expected cumulative reward for taking an action in a given state. The agent refines its policy over time by adjusting these Q-values. The goal of the RL model is to continuously improve its decision-making and respond to new threats in real-time, providing an adaptive solution for automated incident response.

4 RESULTS AND DISCUSSION

4.1 Experimental Results

The RL-based system was tested in a simulated environment and compared with traditional rule-based systems. The results indicated that the RL-based system outperformed traditional methods. Specifically, the RL system detected 95% of threats, while traditional systems detected only 85%. Additionally, the RL system responded 30% faster than rule-based systems, and it had a false positive rate of 3%, compared to 8% for traditional systems, as in Figure 2.

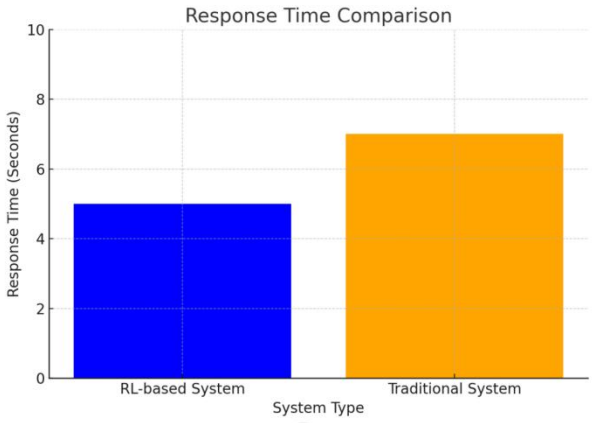
Figure 2 Detection Accuracy Comparison



4.2 Analysis of Results

The RL-based system demonstrated the ability to learn from feedback and adapt to new threats. It consistently improved its decision-making, reducing false positives and optimizing response times over time as in Figure 3.

Figure 3 Response Time Comparison



4.3 Limitations and Challenges

Despite the positive results, several challenges remain. One significant challenge is data dependency. The quality of the training data is crucial for the system's effectiveness, as limited or biased data can affect the RL agent's learning process. Computational complexity is another challenge, as RL models require substantial computational resources, especially for training, which can be a barrier to large-scale deployment. Finally, the RL system may be vulnerable to adversarial attacks that manipulate its learning process.

5 CONCLUSION

This study demonstrates that RL can be highly effective in automating cybersecurity incident response. Traditional rule-based systems, which rely on predefined rules and signatures, often fail to keep up with dynamic and sophisticated cyber threats. In contrast, the RL-based system used in this study showed substantial improvements across key metrics, such as detection accuracy, response time, and false positive rate. Specifically, the RL system achieved a higher detection rate by accurately identifying a wider range of both known and novel threats. Additionally, it outperformed traditional methods in response time, reducing the time needed to mitigate attacks. Another key advantage of the RL system is its ability to reduce false positives, ensuring that security resources are not wasted on false alarms.

The RL agent continuously learns from its feedback, allowing it to adapt to new and evolving cyber threats. This adaptability makes the RL-based system an efficient and scalable solution for real-time incident response. Looking ahead, future research should focus on improving the scalability of RL-based systems to handle large-scale environments. Further work is also needed to enhance the robustness of these models against adversarial attacks that could exploit vulnerabilities in the learning process. Additionally, integrating RL with other AI techniques, such as deep learning, could further improve threat classification and mitigation capabilities.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Alturkistani H, El-Affendi M A. Optimizing cybersecurity incident response decisions using deep reinforcement learning. International Journal of Electrical and Computer Engineering, 2022, 12(6): 6768.

[2] Dunsin D, Ghanem M C, Ouazzane K, et al. Reinforcement learning for an efficient and effective malware investigation during cyber Incident response. arXiv preprint arXiv:2408.01999, 2024.

[3] Ren S, Jin J, Niu G, Liu Y. ARCS: Adaptive Reinforcement Learning Framework for Automated Cybersecurity Incident Response Strategy Optimization. Applied Sciences, 2025, 15(2): 951.

[4] Naseer A, Naseer H, Ahmad A, et al. Moving towards agile cybersecurity incident response: A case study exploring the enabling role of big data analytics-embedded dynamic capabilities. Computers & Security, 2023, 135: 103525.

[5] Manda J K. Cybersecurity Automation in Telecom: Implementing Automation Tools and Technologies to Enhance Cybersecurity Incident Response and Threat Detection in Telecom Operations. Advances in Computer Sciences, 2021, 4(1).

[6] Hassan S K, Ibrahim A. The role of artificial intelligence in cyber security and incident response. International Journal for Electronic Crime Investigation, 2023, 7(2).

[7] Lee Z, Wu Y C, Wang X. Automated Machine Learning in Waste Classification: A Revolutionary Approach to Efficiency and Accuracy. In Proceedings of the 2023 12th International Conference on Computing and Pattern Recognition, 2023: 299-303.

[8] Alturkistani H, El-Affendi M A. Optimizing cybersecurity incident response decisions using deep reinforcement learning. International Journal of Electrical and Computer Engineering, 2022, 12(6): 6768.

[9] Li X, Wang X, Chen X, et al. Unlabeled data selection for active learning in image classification. Scientific Reports, 2024, 14(1): 424.

[10] Liang Y, Wang X, Wu Y C, et al. A study on blockchain sandwich attack strategies based on mechanism design game theory. Electronics, 2023, 12(21): 4417.

[11] Schlette D, Caselli M, Pernul G. A comparative study on cyber threat intelligence: The security incident response perspective. IEEE Communications Surveys & Tutorials, 2021, 23(4): 2525-2556.

[12] Mouratidis H, Islam S, Santos-Olmo A, et al. Modelling language for cyber security incident handling for critical infrastructures. Computers & Security, 2023, 128: 103139.

[13] Oriola O, Adeyemo A B, Papadaki M, et al. A collaborative approach for national cybersecurity incident management. Information & Computer Security, 2021, 29(3): 457-484.

[14] He Y, Zamani E D, Lloyd S, et al. Agile incident response (AIR): Improving the incident response process in healthcare. International Journal of Information Management, 2022, 62: 102435.

[15] Liu Y, Wu Y C, Fu H, et al. Digital intervention in improving the outcomes of mental health among LGBTQ+ youth: a systematic review. Frontiers in psychology, 2023, 14: 1242928.

[16] Wang X, Wu Y C, Ma Z. Blockchain in the courtroom: exploring its evidentiary significance and procedural implications in US judicial processes. Frontiers in Blockchain, 2024, 7: 1306058.
[17] Wang X, Wu Y C, Zhou M, et al. Beyond surveillance: privacy, ethics, and regulations in face recognition technology. Frontiers in big data, 2024, 7: 1337465.
[18] Guo H, Ma Z, Chen X, et al. Generating artistic portraits from face photos with feature disentanglement and reconstruction. Electronics, 2024, 13(5), 955.
[19] Andrade R O, Cordova D, Ortiz-Garcés I, et al. A comprehensive study about cybersecurity incident response capabilities in Ecuador. In Innovation and Research: A Driving Force for Socio-Econo-Technological Development 1st. Springer International Publishing, 2021: 281-29.
[20] Fauziyah F, Wang Z, Joy G. Knowledge Management Strategy for Handling Cyber Attacks in E-Commerce with Computer Security Incident Response Team (CSIRT). Journal of Information Security, 2022, 13(4): 294-311.
[21] Ahmad A, Maynard S B, Desouza K C, et al. How can organizations develop situation awareness for incident response: A case study of management practice. Computers & Security, 2021, 101: 102122.
[22] van der Kleij R, Schraagen J M, Cadet B, et al. Developing decision support for cybersecurity threat and incident managers. Computers & Security, 2022, 113: 102535.