EasyChair Preprint

№ 15665

Reinforcement Learning for Intrusion Detection: Recent Advances and Datasets

Dalal Fathi, Afraa Attiah, Abeer Hakeem and Asmaa Cherif

January 6, 2025

Reinforcement Learning for Intrusion Detection: Recent Advances and Datasets

***Abstract*—With the development of technologies and increasing security threats, intrusion detection systems have become more critical in detecting and protecting operations from attacks. Deep learning has significantly contributed to advancements in intrusion detection, especially through reinforcement learning systems. This survey reviews the concepts of intrusion detection and Reinforcement Learning (RL) systems in intrusion detection, with a focus on recent advancements using techniques such as Multi-Agent Reinforcement Learning (MARL), Adversarial Reinforcement Learning (AE-RL), and Inverse Reinforcement Learning (IRL). We also emphasize the crucial role of feature engineering in conjunction with RL techniques. By examining these cutting-edge approaches and their integration with advanced feature engineering methods, we aim to provide a comprehensive overview of the current state of the art in reinforcement learning-based intrusion detection systems and their potential to enhance cybersecurity measures. In addition to exploring the applications of reinforcement learning and feature engineering in intrusion detection, we highlight and analyze the most well-known databases used in this field, offering insights into the data resources that drive the development and evaluation of these advanced security systems.**

***Index Terms*—Intrusion Detection System (IDS), Reinforcement learning (RL), Feature Engineering, Intrusion Detection Dataset**

I. Introduction

An intrusion detection system (IDS) is a security tool designed to monitor network or system activities for malicious activities or policy violations [1]. It works by scanning network traffic for suspicious patterns that may indicate unauthorized access or attack attempts. This method is required to ensure the correctness, privacy, and accessibility of the system's data. Cybersecurity relies on Intrusion Detection Systems (IDS). However, attackers upgrade their tactics and pose a risk to IDS integrity. Protecting IDS is crucial to safeguard valuable data assets, especially in ongoing digital transformation efforts.

In recent years, the defense against attack techniques, particularly using deep learning, has garnered significant attention [2]. Deep learning algorithms have emerged as powerful tools for enhancing the security framework of network environments through intrusion detection. These advanced algorithms offer enhanced detection capabilities compared to traditional methods for identifying and mitigating evolving threats in real-time [3].

Deep learning has proven successful in various fields due to its ability to handle large-scale data, and researchers have thus focused on its use in intrusion detection [2]. In particular, Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by taking actions in an environment to achieve maximum cumulative reward. It is particularly effective in scenarios where an agent must interact with an environment and learn from feedback through a system of rewards and punishments. Reinforcement learning has been successful in applications such as game playing, robotics, and autonomous systems. In addition, RL offers several advantages for intrusion detection due to its ability to learn from interaction with an environment and make sequential decisions based on feedback. Also, RL agents must balance exploration (trying out new actions) with exploitation (using known actions). In the context of intrusion detection, exploration might lead to attackers exploiting vulnerabilities before effective defenses are learned.

Multi-Agent Reinforcement Learning (MARL) is a subfield of RL that focuses on scenarios where multiple agents interact within a shared environment, and each agent aims to optimize its own objective while considering the presence and actions of other agents [4]. In MARL, there are two or more autonomous agents, each equipped with its own reinforcement learning algorithm, operating within the same environment. These agents may have distinct goals or objectives, and their actions can impact both the environment and the other agents. MARL is particularly relevant in scenarios where multiple autonomous entities need to make decisions while considering the presence and actions of others. This is useful in distributed environments [5].

Adversarial Reinforcement Learning (AE-RL) is a framework that combines elements of reinforcement learning and adversarial training [6]. It involves training a reinforcement learning agent in an environment where it must not only learn to achieve its objectives but also defend against adversarial objects seeking to disrupt its learning process or exploit vulnerabilities in its decision-making. Adversarial reinforcement learning is well-suited for detecting adversarial attacks due to its ability to train agents in adversarial environments, enabling them to learn robust strategies that can withstand and counteract adversarial manipulations.

Recently, a novel approach called Inverse Reinforcement Learning (IRL) has been suggested in [7]. It is a machine learning framework used to infer the underlying reward function of an environment by observing the behavior of an expert agent. Unlike traditional RL, where the reward function is given, inverse reinforcement learning involves recovering the reward function based on an expert's observed behavior.

This work surveys recent advancements in intrusion detection using reinforcement learning techniques, including MARL, AE-RL, and IRL. By examining these cutting-edge approaches, we aim to provide an overview of the current state of the art in reinforcement learning-based intrusion detection systems and their potential to enhance cybersecurity measures. We also emphasize the crucial role of feature engineering in conjunction with RL techniques. Additionally, we presented and analyze the datasets commonly used in conjunction with these reinforcement learning techniques for intrusion detection, offering insights into the data resources that drive the development and evaluation of these advanced security systems.

The remainder of this paper is structured as follows: Section II provides an overview of the research background. Section III discusses the state-of-the-art RL-based IDSs. Section IV reviews the most important datasets used in intrusion detection. Finally, Section V concludes.

II. Research Background

In this section, we define the concept of Intrusion Detection Systems (IDSs), explain reinforcement learning and its subfields, and explore their applications in intrusion detection systems.

A. Intrusion Detection Systems (IDSs)

An IDS is a crucial cybersecurity tool that monitors network or system activities for malicious behavior or policy violations. It primarily detects threats in two ways: signature-based detection and anomaly-based detection [8].

Signature-based detection matches activities against a database of known attack signatures. This method is effective at identifying familiar threats but struggles with new or modified attacks that don't match existing signatures.

Anomaly-based detection, on the other hand, establishes a baseline of normal behavior for the network or system. It then flags any actions that deviate from this baseline as potential threats. This approach is more effective at detecting novel or evolving attack patterns.

B. Reinforcement Learning (RL)

RL is a machine learning technique in which an agent learns to make decisions by interacting with its environment. The agent takes actions, observes the outcomes, and receives reward feedback. The objective is to develop a policy—a mapping from states to actions—that maximizes cumulative rewards over time.

The Key Concepts in Reinforcement Learning are:

- **Agent:** The entity that makes decisions, such as a robot or software.
- **Environment:** The external system with which the agent interacts.
- **State:** A representation of the environment at a specific time.
- **Action:** The choices available to the agent that can change the state.
- **Reward:** Feedback from the environment that indicates the immediate benefit of an action.
- **Policy:** A strategy used by the agent to determine actions based on the current state.

Reinforcement Learning is divided into several subfields based on various criteria. These subfields include methods and strategies used to enhance learning efficiency, adapt to complex environments, and improve performance. The main classifications include:

- Deep Reinforcement Learning (DRL)
- Adversarial Reinforcement Learning (AE-RL)
- Inverse Reinforcement Learning
- Hybrid Reinforcement Learning

1) Deep Reinforcement Learning (DRL): Deep Reinforcement Learning combines reinforcement learning with deep learning techniques, enabling agents to make decisions based on complex, high-dimensional data. In DRL, the agent interacts with its environment and receives feedback—rewards based on its actions. This feedback guides the agent toward better decision-making.

DRL uses neural networks to help an agent decide how to act. These networks estimate two main things: the agent's policy, which guides its actions, and the value function, which predicts expected rewards.

When the agent receives input data, it flows through the neural network. This process helps the agent choose the best actions for different situations. After the agent takes action, it receives rewards that provide important feedback for its improvement.

By utilizing deep neural networks, DRL allows agents to deal with complex tasks with large datasets, including images, audio, and sensor data. Techniques such as Deep Q-Networks (DQN), Policy Gradient methods, and Actor-Critic models make DRL particularly suited for applications in complex decision-making environments, including autonomous driving, robotics, gaming, and intrusion detection.

2) Adversarial Reinforcement Learning: Adversarial Reinforcement Learning (AE-RL) is a framework combining concepts from reinforcement learning (RL) and adversarial training (see Fig. 1). In traditional RL, an agent learns to interact with an environment to maximize its cumulative reward signal by taking actions that influence the state transitions and the rewards it receives.

AE-RL involves two agents: an agent classifier and an agent environment [9]. The agent classifier identifies and classifies network intrusions and anomalous activities. The agent environment simulates the network environment. The

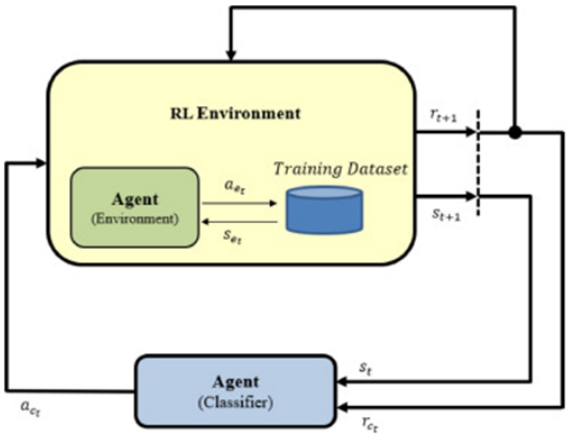

Fig. 1. Adversarial Reinforcement Learning [6]

environment dynamics include new network packets, adversarial attacks, and the consequences of the agent's actions [6]. By modeling the network environment adversarially, the agent environment challenges the agent classifier to adapt and improve its detection capabilities against evolving threats and attack strategies.

3) Inverse Reinforcement Learning (IRL): RL and IRL offer contrasting approaches to machine learning (see Fig. 2).

RL is a specific branch of machine learning that emphasizes training an agent to make decisions based on interactions with an environment. The agent learns how to make better decisions through trial and error by receiving feedback in the form of rewards or punishments from the environment. It finds the best actions to take in a given situation based on this feedback. In this approach, an agent interacts with its environment and adjusts its behavior to maximize the total reward it receives.

On the other hand, IRL is a technique that infers the underlying reward function from observed behavior to understand decision-maker goals. Rather than assuming a fixed reward function, inverse reinforcement learning tries to infer the reward function that would explain the behavior of an expert in a given task [10].

(a) RL (b) IRL

Fig. 2. Reinforcement Learning vs Inverse Reinforcement Learning

By inferring the reward function, IRL enables the agent to learn from the expert's behavior and understand their goals, even if they are not explicitly stated. This can be especially useful when it is challenging to specify a clear reward function, such as in complex real-world applications.

IRL is a different approach to Imitation Learning (IL) [11]. IL is effective for copying expert behavior, while IRL is more useful for understanding the reasons behind their actions, enabling us to adapt in various situations.

4) Hybrid Reinforcement Learning:: Hybrid reinforcement learning models, which combine traditional reinforcement learning techniques with other machine learning approaches, have shown promising results in various domains [12, 13]. These models leverage the strengths of different algorithms to overcome limitations of pure reinforcement learning, such as sample inefficiency and slow convergence [14]. Hybrid RL models that combine two different RL approaches have gained significant attention in recent research. These models aim to leverage the strengths of multiple RL techniques to overcome limitations and improve performance in complex decision-making tasks. For instance, [15] combines adversarial training with IRL to learn robust reward functions.

C. Feature Engineering

Enhancing data representation in machine learning models requires feature engineering. The main techniques used in feature engineering are feature transformation and feature selection. Feature selection identifies essential data components, while feature transformation optimizes their representation.

1) Feature Transformation: Feature transformation modifies or creates new features from original data to aid pattern recognition. Key methods include:

- Scaling and normalization: Makes features like packet sizes or timestamps comparable.
- Encoding categorical data: Converts protocol types into numerical values.

2) Feature Selection: Feature selection involves identifying and retaining only the most impactful features in a dataset. This process lowers dimensionality and computational requirements, allowing the model to concentrate on data that most significantly influences its learning. Important techniques include:

- Filter methods: Such as correlation tests
- Wrapper methods: Like Recursive Feature Elimination
- Embedded methods: Apply model-specific criteria, such as regularization in Lasso regression

Feature selection can be implemented using automated methods, such as filter and wrapper techniques, or through embedded methods, where the learning algorithms, including RL algorithms, dynamically select features during the learning process (see Fig. 3). This integration is particularly useful in complex environments like robotics, gaming, and intrusion detection, where the state space is high-dimensional, and efficient learning is crucial for optimal performance.

In IDS applications, feature selection might prioritize core traffic attributes—such as packet count, flow duration, and port access patterns—while discarding features that do not improve the model's ability to differentiate normal from malicious traffic. Feature selection identifies essential data components, while feature transformation optimizes their representation. In RL-based IDSs, these processes work together—feature selection reduces noise, and transformation enhances the features, allowing the RL model to more effectively understand and respond to threats.

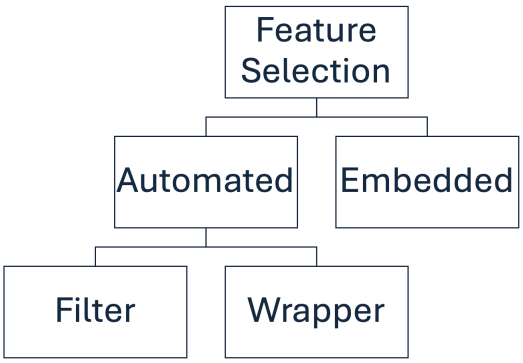

Fig. 3. Feature Selection classification

III. Application of Reinforcement Learning in IDSs

This section reviews significant research works according to the taxonomy illustrated in Fig. 4. We categorize the literature based on RL approaches in IDS, including traditional RL, IRL, and hybrid models.

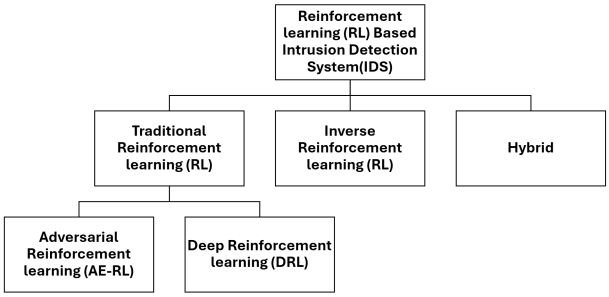

Fig. 4. RL-based IDS Taxonomy

A. Traditional Reinforcement Learning (RL)

1) AE-RL-based Models: The models presented in [4], [6] [9] and [16] employed a two-agent system for network intrusion detection: an agent classifier and an agent environment. These components work in tandem to create a robust and adaptive IDS.

The agent classifier serves as the core of the IDS, focusing on accurately identifying and classifying network intrusions or anomalous activities within network traffic data. It learns to differentiate between normal behavior and potential security threats by analyzing features extracted from network packets or flows. This capability enables the system to make informed decisions and facilitate effective real-time detection and mitigation of security risks.

Complementing the classifier, the agent environment simulates the network ecosystem in which the classifier operates. It models the interactions between the agent and network traffic data, providing crucial feedback to the classifier based on its decisions. The environment's dynamics include new network packet arrivals, potential adversarial attacks, and the consequences of the agent's actions. By incorporating an adversarial setting, the agent environment challenges the classifier to continuously adapt and improve its detection capabilities against evolving threats.

Mouyart et al. [4] addressed data imbalance and bias by using a generative model to augment the dataset, improving its robustness. The DRL approach allows each agent to independently analyze and learn from network patterns, while the generative model ensures balanced data representation.

Guillermo et al. [6] proposed an algorithm that formulates intrusion detection as a Markov Decision Process (MDP Q-Learning). It is designed to adapt dynamically to evolving attack strategies by leveraging RL in an adversarial environment.

Suwannalai et al. [9] leveraged AE-RL with Deep Q-Networks (DQN) to develop a robust and adaptive IDS. This approach focused on training the IDS using a deep reinforcement learning algorithm, enabling it to learn an optimal policy for detecting network intrusions through continuous interaction with the environment and reward-based feedback.

Mahjoub et al. [16] proposed an AE-RL algorithm for innovative intrusion detection in IoT systems. The algorithm integrates supervised and adversarial RL models, creating a simulation environment adhering to RL principles. This pioneering study applies AE-RL to IoT intrusion detection, introducing a novel architecture that combines supervised and adversarial RL models to address prediction challenges in demanding IoT networks.

2) DRL-based Models: Hsu et al.[17] utilized the NSL-KDD dataset with a DRL-based anomaly detection model, achieving high accuracy in identifying novel attacks. Their DRL anomaly detection engine features two modes: detection and learning, which can be flexibly switched. The system switches to learning mode whenever the detection mode's performance falls below a predefined threshold, allowing it to learn new network traffic patterns. They evaluated their approach using two well-established benchmark network intrusion simulation datasets: NSL-KDD and UNSW-NB15. Additionally, they applied their method to their campus network environment, which comprises approximately 300 million daily network traffic records, about 100 times larger than either of the synthetic datasets.

Manuel et al. [18] applied DRL within a supervised framework using a custom-labeled dataset, resulting in improved accuracy and response rates for complex, real-time detection scenarios. The authors modified the classic DRL paradigm by replacing the live environment with a sampling function of recorded training intrusions. This new pseudo-environment samples the training dataset and generates rewards based on detection errors during training. The technique was applied to four DRL models: DQN, DDQN, PG, and AC, with DDQN yielding the best results. The study provides a comprehensive comparison of results using the AWID and NSL-KDD datasets.

Alavizadeh et al. [19] developed a network intrusion detection method that combines Q-learning reinforcement learning

with deep feed-forward neural networks. The study details the fine-tuning of hyperparameters for optimized performance. Experimental results using the NSL-KDD dataset demonstrate the method's high effectiveness in detecting various intrusion classes, outperforming similar machine learning approaches with over 90% accuracy in classifying different network intrusion types. This advancement represents a significant improvement in network security, offering enhanced adaptability and accuracy in intrusion detection.

[20] introduced a DRL model to enhance feature selection for IDS. This model combined recursive feature elimination (RFE) with DRL, improving the identification of key features for classification tasks. The main innovation was a hybrid approach that merged RFE's efficiency in feature selection with DRL's performance-boosting abilities. Testing on the CSE-CIC-IDS2018 dataset showed impressive results, reducing redundant features by 80% and enhancing detection accuracy in complex networks. While RFE offered benefits like improved accuracy and lower computational requirements by removing irrelevant features, it had some drawbacks. These included being time-consuming due to repeated model fitting and dependence on the underlying model's performance, which could limit its applicability to other models.

B. IRL-based Models

Lian et al. [7] developed a method for optimizing autonomous systems using Inverse Reinforcement Learning (IRL) to detect and correct anomalies. In this approach, an agent is trained to learn the optimal behavior of the autonomous system from expert demonstrations or data. The study's strength lies in developing an effective anomaly detection and correction method for optimizing autonomous systems using inverse reinforcement learning, as evidenced by successful verification through simulations and experiments on a quadrotor UAV.

Parras et al. [21] proposed defense strategies that use IRL to detect smart attackers in wireless networks. The experimental setup involved testing these defense mechanisms against an intelligent attacker who employs DRL in a back-off attack scenario. The results demonstrated that the IRL-based defense mechanisms effectively identified intelligent attackers and performed well even under conditions of partial observability.

Fan et al. [22] developed a new framework for feature selection that balanced effectiveness and efficiency. Their main contribution was creating an interactive reinforcement learning (IRL) model that used decision tree feedback to improve feature selection. This framework enhanced feature representation by combining feature-feature graphs, decision tree structures, and personalized reward systems. Results showed this approach performed better than traditional feature selection methods, especially with complex datasets. It employed a feedback loop between the IRL and decision tree to refine feature importance and decision-making, outperforming existing methods. The proposed approach effectively found optimal feature subsets with better accuracy and efficiency, utilizing interactive learning and structured feedback from decision trees. However, the model's complexity could lead to higher computational costs, particularly in managing interactions between agents and external trainers, and might require careful tuning of hyperparameters. The paper suggested exploring other downstream tasks beyond decision trees and investigating different types of trainers for the interactive loop, which could further improve feature selection across various datasets and applications.

C. Hybrid Models

Several studies have explored hybrid models combining different RL techniques or integrating RL with other machine-learning approaches

Najafli et al. [23] and Alhaddad et al. [24] propose hybrid Deep Q-learning models combined with Gated Recurrent Units (GRU).

[23] focused on enhancing security in fog-to-cloud computing environments. The model features a two-layer architecture: the fog layer conducts binary classification to identify normal or malicious traffic, while the cloud layer performs multi-class classification for specific attack types. This design optimizes resource use and improves accuracy in detecting cyber threats. The IDS enhances detection accuracy and reduces false negatives by combining DQL with GRU for temporal analysis and ensemble methods. Evaluated using the CIC-IDS2018 dataset, future work suggests using additional datasets, addressing data imbalance, and optimizing RL parameters to boost the system's effectiveness in real-world IoT environments.

[24] targeted the detection of Distributed Denial-of-Service (DDoS) attacks within Smart Grid environments. The study's significant contributions include developing hybrid models and installing a real-time monitoring dashboard for live threat identification and visualization. Tested on the CIC-DDoS2019 dataset and a custom dataset, the model achieved a detection accuracy of 99.86%, with precision and F1 scores close to 99.5% and 99.68%, respectively, marking a significant improvement over existing detection methods.

Strickland et al. [25] presented a model combining DRL and Generative Adversarial Networks (GAN) for enhanced network IDS. It uses the NSL-KDD and CICIDS2017 datasets to perform binary and multiclass classification efficiently, improving detection accuracy and durability against adversarial attacks. Despite enhanced detection rates and adaptability, the DRL-GAN model faces challenges due to the training complexity and resource requirements of GANs. Future studies will focus on reducing computational complexity and enhancing adaptation to new threats to increase scalability in real-time IDS systems.

Liu et al. [26] presented a novel multi-agent reinforcement learning system for feature selection. Their main contribution was recasting the feature selection problem by treating each feature as an agent, allowing for a more thorough and efficient exploration of feature subsets. The system utilized autoencoders, graph convolutional networks (GCN), and meta-descriptive statistics for state representation, along with a generative rectified sampling strategy to boost training efficiency.

Extensive tests demonstrated that this method significantly outperformed traditional approaches like K-Best Selection, LASSO, and genetic algorithms in prediction accuracy and efficiency. This framework enhanced cooperation and competition among feature agents, leading to better feature subset selection. The benefits of this approach included high prediction accuracy while maintaining flexibility and adaptability across various datasets. The multi-agent system ensured a comprehensive exploration of feature spaces, resulting in improved performance in feature selection tasks. However, there were some drawbacks, including the model's complexity, which could increase computational costs and require extensive fine-tuning. Additionally, relying on deep learning-based state representations might cause training issues and longer convergence times. The authors suggested exploring more ways to improve exploration efficiency and adapting the framework for real-time feature selection in changing environments. Future research could also focus on integrating other reinforcement learning methods and enhancing state representations to improve the approach's robustness.

Table I summarizes the application of RL in IDS. It highlights the RL method used, the feature selection category, the methodology, the learning algorithm, and the best performance. We observe that only a few research investigated IRL models and sophisticated feature selection methods.

IV. Intrusion Detection Dataset used with RL

A crucial element in advancing research on Intrusion Detection Systems (IDS) using Reinforcement Learning (RL) is the availability of well-defined datasets. These datasets serve as a foundation for training, validating, and testing RL algorithms within IDS frameworks, enabling accurate and robust threat detection. Table II summarizes the key datasets commonly used in IDS.

The NSL-KDD dataset [27] comprises a total of 125,973 records, with 25,973 reserved for testing. By removing duplicate items, this dataset enhances the KDD'99 dataset, allowing for a more impartial assessment of intrusion detection systems.

The AWID dataset [28] includes tagged recordings from a wireless network and focuses on wireless intrusion detection. It consists of 545.730 records and 151,102 records in the testing set, helping researchers assess detection capabilities in this specific domain.

The CSE-CIC-IDS2018 dataset [29] is significantly larger, featuring 15,000,000 records. 10,000,000 record are set aside for testing. This dataset evaluates the effectiveness of intrusion detection systems by simulating a variety of attacks in realistic environments.

The BOT-IoT dataset [30], which targeted botnet attacks within IoT networks, contains approximately 1,083,000 records. This dataset used 78,000 record for testing, providing a comprehensive view of IoT security challenges.

The CICIDS2017 dataset [31] includes 2,580,000 records, with 1,230,000 for testing. This dataset simulates real-world attack scenarios alongside normal network traffic, which aids in the assessment of intrusion detection techniques.

Lastly, the CIC-DDoS2019 dataset [32] features 5,800,000 records, and 1,100,000 are designated for testing. This dataset was created specifically to replicate modern DDoS attack scenarios, assisting in the development of efficient detection and mitigation techniques.

V. Conclusion

This survey offers a comprehensive overview of intrusion detection, emphasizing its fundamental concepts and various types, including signature-based detection and anomaly-based detection. Signature-based detection relies on known patterns of malicious activity, while anomaly-based detection identifies deviations from normal behavior. Additionally, the survey delves into the application of reinforcement learning in the realm of intrusion detection, showcasing how this advanced technique can enhance detection capabilities.

Moreover, the survey underscores the importance of selecting the right features during the detection process and classifie spapers accordingly. We also examine the different types of datasets utilized in this field, which are crucial for training and evaluating detection models effectively. The survey illustrates how these interconnected concepts play a significant role in fortifying cybersecurity measures. Our investigation reveals there is a lack in research concerning the use of inverse reinforcement learning and advanced feature selection methods. Addressing this gap could pave the way for future research directions, leading to the development of more precise and effective solutions in the field of intrusion detection.

References

[1] Ansam Khraisat et al. "Survey of intrusion detection systems: techniques, datasets and challenges". In: *Cybersecurity* 2.1 (2019), pp. 1–22.

[2] Arwa Aldweesh, Abdelouahid Derhab, and Ahmed Z. Emam. "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues". In: *Knowledge-Based Systems* 189 (2020), p. 105124. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2019.105124. URL: https://www.sciencedirect.com/science/article/pii/S0950705119304897.

[3] Priyanka Dixit and Sanjay Silakari. "Deep Learning Algorithms for Cybersecurity Applications: A Technological and Status Review". In: *Computer Science Review* 39 (2021), p. 100317. ISSN: 1574-0137. DOI: https://doi.org/10.1016/j.cosrev.2020.100317. URL: https://www.sciencedirect.com/science/article/pii/S1574013720304172.

[4] Matthieu Mouyart, Guilherme Medeiros Machado, and Jae-Yun Jun. "A multi-agent intrusion detection system optimized by a deep reinforcement learning approach with a dataset enlarged using a generative model to reduce the bias effect". en. In: *J. Sens. Actuator Netw.* 12.5 (Sept. 2023), p. 68.

TABLE I
SUMMARY OF REINFORCEMENT LEARNING-BASED TECHNIQUES IN IDSS

Ref#	**RL method**	**Feature Engineering**	**Method Selection**	*Alogrithm*	*Accuracy*	*F1 Score*
[22]	IRL	Selection	Decision Tree-Based Trainer, KBest Based Trainer	DRL + RFE	96.18	94.89
[20]	DRL	Selection (Automated)	Reinforced Feature Elimination with Deep Reinforcement Learning	DRL + RFE	96.18	94.89
[26]	Hybrid	Selection (Embedded)	Multi-agent Reinforcement Learning Feature Selection	MARLFS	95.6	94.92
[9]	AE-RL	Transformation	Adversarial Deep Q-Learning Network	AE-DQN	80	79
[6]	AE-RL	Transformation	Reinforcement Learning with Adversarial Objectives	AE-RL	80	79
[16]	AE-RL	Transformation	Adversarial Reinforcement Learning with Supervised Learning	AE-DQN	99.98	94.59
[19]	DRL	Transformation	Deep Q-Learning for anomaly classification	DQL	90	88
[18]	DRL	Transformation	Double Deep Q-Learning Network (DDQN) for supervised intrusion detection	DDQN	89.5	87.5
[7]	IRL	Transformation	Inverse Reinforcement Learning for anomaly correction in autonomous systems	IRL	85	84.5
[21]	IRL	Transformation	Inverse Reinforcement Learning for attack mitigation in network backoff	IRL	88	86
[23]	Hybrid	Transformation	Deep Reinforcement Learning and Ensemble Method	DQL+GRU	99.99 for Botnet attacks	99.59
[25]	Hybrid	Transformation	Deep Reinforcement Learning combined with GAN	DRL+GAN	89.5	73.0
[24]	Hybrid	Transformation	Hybrid of CNN and GRU	DL(CNN)+GRU	99.86.	Near 100 for DDoS
[17]	DRL	Transformation	Deep Reinforcement Learning for anomaly detection	DRL	85	82

TABLE II
OVERVIEW OF DATASETS USED IN REINFORCEMENT LEARNING INTRUSION DETECTION SYSTEMS

Dataset	*References on RL in IDS*	*Total Record*	*Testing Record*	*Focus Area*
NSL-KDD	[6][9][19][18] [25]	125,973	25,973	Intrusion detection system evaluation
AWID	[18]	545,730	151,102	Wireless intrusion detection
CSE-CIC-IDS2018	[23]	15,000,000	10,000,000	Simulation of various attack types in realistic settings
BOT-IoT	[16]	1,083,000	78,000	Botnet attacks in IoT networks
CICIDS2017	[25]	2,580,000	1,230,000	Real-world attack simulation and normal traffic
CIC-DDoS2019	[24]	5,800,000	1,100,000	DDoS attack scenarios

[5] Mohamed Amine Ferrag et al. "Deep Learning-Based Intrusion Detection for Distributed Denial of Service Attack in Agriculture 4.0". In: *Electronics* 10.11 (2021). ISSN: 2079-9292. DOI: 10.3390/electronics10111257. URL: https://www.mdpi.com/2079-9292/10/11/1257.

[6] Guillermo Caminero, Manuel Lopez-Martin, and Belen Carro. "Adversarial environment reinforcement learning algorithm for intrusion detection". In: *Computer Networks* 159 (2019), pp. 96–109. ISSN: 1389-1286. DOI: https://doi.org/10.1016/j.comnet.2019.05.013. URL: https://www.sciencedirect.com/science/article/pii/S1389128618311216.

[7] Bosen Lian et al. "Anomaly Detection and Correction of Optimizing Autonomous Systems With Inverse Reinforcement Learning". In: *IEEE Transactions on Cybernetics* 53.7 (2023), pp. 4555–4566. DOI: 10.1109/TCYB.2022.3213526.

[8] Zhuowei Li, Amitabha Das, and Jianying Zhou. "Usaid: Unifying signature-based and anomaly-based intrusion detection". In: *Advances in Knowledge Discovery and Data Mining: 9th Pacific-Asia Conference, PAKDD 2005, Hanoi, Vietnam, May 18-20, 2005. Proceedings 9*. Springer. 2005, pp. 702–712.

[9] Ekachai Suwannalai and Chantri Polprasert. "Network Intrusion Detection Systems Using Adversarial Reinforcement Learning with Deep Q-network". In: *2020 18th International Conference on ICT and Knowledge Engineering (ICT&KE)*. 2020, pp. 1–7. DOI: 10.1109/ICTKE50349.2020.9289884.

[10] Alexander Peysakhovich. "Reinforcement learning and inverse reinforcement learning with system 1 and system 2". In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 2019, pp. 409–415.

[11] Bilal Piot, Matthieu Geist, and Olivier Pietquin. "Bridging the Gap Between Imitation Learning and Inverse Reinforcement Learning". In: *IEEE Transactions on Neural Networks and Learning Systems* 28.8 (2017), pp. 1814–1826. DOI: 10.1109/TNNLS.2016.2543000.

[12] Zongzhang Wang et al. "Hybrid reinforcement learning: A comprehensive survey". In: *arXiv preprint arXiv:2009.06862* (2020).

[13] Qiang Li et al. "Hybrid actor-critic reinforcement learning in parameterized action space". In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 2019, pp. 3232–3238.

[14] Kai Zhang, Zhuoran Yang, and Tamer Basar. "A survey of deep reinforcement learning in video games". In: *arXiv preprint arXiv:1912.10944* (2021).

[15] Justin Fu, Katie Luo, and Sergey Levine. "Learning robust rewards with adversarial inverse reinforcement learning". In: *arXiv preprint arXiv:1710.11248* (2018).

[16] Chahira Mahjoub et al. "An adversarial environment reinforcement learning-driven intrusion detection algorithm for Internet of Things". In: *EURASIP Journal on Wireless Communications and Networking* 2024.1 (May 2024), p. 21. ISSN: 1687-1499. DOI: 10.1186/s13638-024-02348-6. URL: https://doi.org/10.1186/s13638-024-02348-6.

[17] Ying-Feng Hsu and Morito Matsuoka. "A Deep Reinforcement Learning Approach for Anomaly Network Intrusion Detection System". In: *2020 IEEE 9th International Conference on Cloud Networking (CloudNet)*. 2020, pp. 1–6. DOI: 10.1109/CloudNet51028.2020.9335796.

[18] Manuel Lopez-Martin, Belen Carro, and Antonio Sanchez-Esguevillas. "Application of deep reinforcement learning to intrusion detection for supervised problems". In: *Expert Systems with Applications* 141 (2020), p. 112963. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2019.112963. URL: https://www.sciencedirect.com/science/article/pii/S0957417419306815.

[19] Hooman Alavizadeh, Hootan Alavizadeh, and Julian Jang-Jaccard. "Deep Q-Learning Based Reinforcement Learning Approach for Network Intrusion Detection". In: *Computers* 11.3 (2022). ISSN: 2073-431X. DOI: 10.3390/computers11030041. URL: https://www.mdpi.com/2073-431X/11/3/41.

[20] Kezhou Ren et al. "ID-RDRL: a deep reinforcement learning-based feature selection intrusion detection model". In: *Scientific reports* 12.1 (2022), p. 15370.

[21] Juan Parras et al. "Inverse Reinforcement Learning: A New Framework to Mitigate an Intelligent Backoff Attack". In: *IEEE Internet of Things Journal* 9.24 (2022), pp. 24790–24799. DOI: 10.1109/JIOT.2022.3194694.

[22] Wei Fan et al. "Interactive reinforcement learning for feature selection with decision tree in the loop". In: *IEEE Transactions on Knowledge and Data Engineering* 35.2 (2021), pp. 1624–1636.

[23] Sepide Najafli, Abolfazl Toroghi Haghighat, and Babak Karasfi. "A novel reinforcement learning-based hybrid intrusion detection system on fog-to-cloud computing". In: *The Journal of Supercomputing* 80.18 (2024), pp. 26088–26110.

[24] Ulaa AlHaddad et al. "Ensemble model based on hybrid deep learning for intrusion detection in smart grid networks". In: *Sensors* 23.17 (2023), p. 7464.

[25] C Strickland, C Saha, M Zakar, et al. "DRL-GAN: A Hybrid Approach for Binary and Multiclass Network Intrusion Detection. arXiv preprint arXiv: 230103368". In: (2023).

[26] Kunpeng Liu et al. "Automated feature selection: A reinforcement learning perspective". In: *IEEE Transactions on Knowledge and Data Engineering* 35.3 (2021), pp. 2272–2284.

[27] M Hassen ZAIB. *NSL KDD Dataset*. https://www.kaggle.com/datasets/hassan06/nslkdd. [Online; accessed 19-July-2024]. 2024.

[28] Zhiqing Cui. *AWID CLS-R Dataset*. Accessed: 2024-10-31. 2020. URL: https://www.kaggle.com/datasets/zhiqingcui/awidclsr.

[29] Solarmainframe. *CSE-CIC-IDS2018*. https://www.kaggle.com/datasets/solarmainframe/ids-intrusion-csv. Accessed: 2024-10-31. 2023.

[30] Vignesh Venkateswaran. *Bot-IoT Dataset*. Accessed: 2024-10-31. 2019. URL: https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot.

[31] Koo Aslansefat. *CICIDS2017 SafeML Notebook*. https://www.kaggle.com/code/kooaslansefat/cicids2017-safeml. Accessed: 2024-10-31. 2022. URL: https://www.kaggle.com/code/kooaslansefat/cicids2017-safeml.

[32] d. Hoogla. *CIC-DDoS2019 Dataset*. Accessed: 2024-10-31. 2022. URL: https://www.kaggle.com/datasets/dhoogla/cicddos2019.