Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

A Deep Reinforcement-Based Anomaly Intrusion Detection for Enhancing Network Cybersecurity

Maytham Mohammed Tuaama
Imam Al-Kadhum College (IKC), Department of Computer Technical Engineering.
maythammohammed@iku.edu.iq

Abstract.

Conventional protection methods, such as rules-based firewalls and signature-based detection, are not cutting it in today's environment of increasingly sophisticated and frequent cyberattacks. Cyberattacks nowadays are extremely dynamic and complex, calling for cutting-edge solutions that can change and adapt as the threat does. DRL is an AI subfield that has been successfully addressing difficult decision-making challenges in several fields, including cybersecurity. Here, we make a step forward by using a DRL framework to model cyberattacks; by incorporating real-world events, we make the models more realistic and applicable. We provide a customized approach that greatly improves existing approaches by carefully tailoring DRL (deep reinforcement algorithms to the complex needs of cybersecurity situations, including adversarial training, dynamic environments, bespoke structure of reward and actions, and more.
In this study, we provide an anomaly detection method to detect attacks on network CPS using Deep Reinforcement Learning. Our proposed methodology was tested using several publicly available research datasets to ensure its efficacy.

Keywords: - ML; DRL; CPS; Reinforcement Learning; Cyber Physical System; DL

1. Introduction

In an era of greater network connectivity, cybersecurity has become an urgent issue, particularly when addressing the utmost goal of network security protection [1]. While several anomaly detection models are widely used for revealing unknown types of attacks, numerous intrusion detection systems (IDSs) have not achieved satisfactory intrusion detection performance [2]. This is because the recent development of rapport, behavior, and attack correlations over network traffic data involves high complexity that requires advanced model construction[3] . The rising popularity of network-based intrusion detection has highlighted many limitations that need to be addressed as more advanced evasion attack techniques become available. As a result, identifying an efficient anomaly intrusion detection system has tremendous potential to greatly impact cybersecurity protection [4].

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

The recent establishment of deep learning principles for various classification tasks has constituted one of the largest technology shifts in recent years[5] . When considering their design principles, these neural network architectures - trained by vast amounts of labeled data - have demonstrated model capabilities on diverse applications [6]. With the surge of deep learning in numerous areas, the pattern recognition community has extended it to drive significant advances in IDS. As instituted by the deep belief network era, researchers have developed IDS models with deep learning techniques and have demonstrated improved detection capability compared to cutting-edge methods. Those proposed deep neural network models outperform previous models and operate much faster than more complicated models, even while being trained with large datasets [7]. Given the importance and demand for enhanced IDSs, applying various deep learning algorithms to IDS models will continue, and research in this field will provide further benefits.

1.1. Background and Motivation

The network and the cyberinfrastructure are evolving continuously at an incredible speed. The number of IoT devices is constantly growing. At the same time, with the evolution of the mobile cloud, various new applications that deal with large volumes of data have been introduced[8], [9]. Moreover, many applications have been developed in multifaceted domains to enable factors. Their mission is to solve many issues in areas such as environmental protection, better urban management, better transportation solutions, better healthcare solutions, and a more secure society [10]. Trusting as we can be these days in the more advanced technologies that support a big community as urban areas, we all know that the world is dangerous. There are many different harmful aspects that people can and want to unleash [11], [12].

Zero-day attacks target network vulnerabilities unknown to information security[13]. The firewall and IDS are the two mainstream network security technologies. Still, they are both signature-based detection techniques, which are essentially historical judgement systems: usually unable to detect zero-day attacks in advance[14], [15]. Conversely, zero-day attacks aim to exploit known or previously unknown network vulnerabilities, trying to intrude into networks to cause damage or obtain some secret information[16]. The zero-day exploit exhibits a self-adapting behavior, as it could be stable or fluctuating. With the increasing complexity and diversity of zero-day attacks, signature-based techniques have become less efficient and effective in their detection/ prevention because of the manual work (or not) involved in the process, the higher workload in real-life networks, the occurrence of false positives, and the higher computational costs related to the continuous signature update[17]. The anomaly, artificial intelligence-based IDS, addresses the limitations of signature-based detection by learning and evolving, leading to—the planning of—advanced attacks. The non-availability of large-labeled data, their high imbalance, and the susceptibility of a trained model to adversarial attacks are some of the problems with the data-driven anomaly detection methods [18], [19], however. In this paper, we present a deep reinforcement-based network anomaly intrusion detection (DRNAID) model, which is the result of our work presented in this paper.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

The DRNAID model performs anomaly detection mostly via two sub-models. The first presented sub-model is a pre-training step intended to speed up the convergence of the other sub-model: the DRNAID-RTD model. The previously trained model creates, by effective and accurate ability, detecting the complex volumetric attacks, an environment map that the other model can learn on. Each training model learns through simulations based on the respective environment toward optimizing its learning goal. Then, the training model gives feedback on its learning result through SGD algorithms by comparing the predicted output and real value input. Adjusting the reinforced mechanism involves AI for improved health and increases its prediction accuracy at the end of each process, enhancing the AI's intelligence. The methodology introduced in the article conducted an experimental study where the results showed that the hybrid model presented an improved prediction accuracy as compared to traditional models based on the validation of the experimental results with a t-test.

1.2. Research Objectives

By learning from the successes of human specialists who work in these security operations centers, the main idea of this paper is to leverage these intelligent data-driven models, especially deep reinforcement learning, and then design a network security management agent for enhancing computer network security monitoring. Our research carefully applies the actor-critical model architecture in designing the security reinforcement learning agent, considering the characteristics and requirements of network traffic when it over time. The actor-critic model combines policy-based (actor) and value-based (critic) into a single model. The actor recommends network security decisions (policy) given the state. The critic is responsible for evaluating the actor's decision (value) given that state in real-time.

Consequently, the agent will learn from cumulative internal numerical feedback called reward, obtained by predicting the attack level for some samples and feedback to the model. Upon learning from the attacker's reward pattern, the agent seeks spectrum in the traffic spectrums. By getting hands-on experiences from the simulated intrusion attacks, the agent will continuously and incrementally modify its knowledge (policy) on defensive resource allocations. In short, after training, the reinforcement learning agent is expected to automatically and efficiently configure security monitoring settings (on-the-fly decision) that optimize the detection performance of anomaly-based intrusion detection algorithms.

2. Anomaly Detection in Network Security

Anomaly detection is a technique used to identify unusual patterns that do not match the common behavior. The anomaly detection technique can be classified into three types based on supervised, unsupervised, and semi-supervised learning [20]. In the case of supervised learning, the training process can consume an enormous amount of time as the class labels in the training data need to be accurate and fully representative [21]. Required labor costs and a significant volume of labeled data are two compelling problems when developing anomaly detection models [22]. The second and third techniques can greatly help to solve these problems. The unsupervised anomaly detection approach requires no class labels, so the training dataset can be constructed

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

more easily [23]. Despite this advantage, unsupervised techniques face higher class imbalance ratios in the network traffic data[24]. Furthermore, the pure network environment only has normal data. At the same time, no anomaly is labeled to the intrusion detection system, so the network traffic data can potentially contain many uncommon usages that are helpful to the attacker in reaching their targets[25].

Some researchers have focused on anomaly detection-related intelligence techniques in the past few years, including deep learning for the network traffic dataset. Deep learning is a novel, more potent technique than conventional machine learning[26]. Compared to them, it has more data, more complicated models, and acceptable computational costs [27]. Furthermore, deep learning can consume old features or multiple feature representations to effectively hold onto non-linear correlations between features [28]. On the other hand, based on the well-publicized NSL-KDD dataset and considered more effective than others, the Network Intrusion Detection Technique achieves higher performance. In practice, not only were the above-mentioned effective results questioned, but deep-based methods for dealing with the NSL-KDD dataset also challenged the ability of the network security dataset [29]. They did not achieve satisfactory results because the latter contains information useful to detect many network traffic anomalies, including the intrusion-created port connections used to scan the local network and the doorknob extracted from the NSL-KDD dataset, among others [30].

2.1. Types of Anomalies

Two major types of anomalies can occur when comparing the behavior of systems or users on networks: Type I and Type II anomalies. Type I anomalies argue with departure from normal behavior in network connections, while Type II anomalies point to point-to-point connections[31] . Type I anomalies are characterized mainly by deviation from statistical properties such as PDF, spectral radius, statistical moment, and SVD singular values. Simple statistical analysis allows effective detection of Type I anomalies[32]. On the other hand, if the characteristics of PDF and RR properties in NS and P2P flow are extracted, however, Sansinene et al. claim that Type II anomalies can also be detected effectively[33].

In the anomaly detection framework associated with the available analysis results on spatial features in a linear associative memory model of a Gaussian nature, FANNG generates 'S-Point' as a function of the adaptive correlation coefficient and uses it for anomaly detection. Sansinene et al. note that Anomaly differs from network behavior that tends to exceed normal region bandwidth because 'S-Point' can guarantee a continuous flow of standard peak traffic [30].

2.2. Challenges in Anomaly Detection

For designing the anomaly detection model for a rugged network, research should consider some inherent characteristics of network traffic data[34] . The key challenge here is the explosion in size, both in terms of quantity and dimensionality of network traffic data. Four main challenges are associated with detecting a network's abnormal traffic grade: data imbalance, varying patterns, feature ambiguity, and multi granularity

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[35]. According to the risk probability, the number of abnormal data points in a network traffic dataset is in the minority classes. Occasional consistency problems occur during the training time when learning models by optimizing common classification measures, such as the total sum of all parts of squared errors, for the plethora of normal data compared to their minimal abnormal counterparts [36]. For unbalanced issues, the model pertained can be skewed, alerting about potential issues with data bias. As traffic data is continuously produced, it changes dynamically and intermittently owing to the network's large and disparate topology. Due to adaptive sampling, data mining methods based on sliding a generally structured approach like a moving window may succinctly elucidate the temporary dynamics of a stream and change in knowledge over time [24]. However, with widely varying data representations like calendar-based features and short-term, frequent variation-based features, classic data processing techniques apply only to a fixed number of features in spacetime [37].

3. Reinforcement Learning in Cybersecurity

Reinforcement learning (RL) is about agents observing their states and acting in a series of moments to take decisions whose effects will be observed in the future [38]. They aim to maximize some overall goal by collecting reinforcements, or rewards, on each decision. It is also the best-tested way to process large interdependent datasets that arrive in multiple varieties to develop understanding and make fair, efficient choices using that understanding[39]. This broad ability makes reinforcement learning particularly qualified to investigate and address numerous significant cybersecurity issues [40].

An intelligent, autonomous, and powerful cyber network contains large numbers of interconnected cyber devices and applications. These devices and applications can have different features and function differently. For instance, data and control plane protocols are commonly used in these cyber-network devices [41]. Such functions may cause network attack states and violate network security. As such, intrusions can become extremely difficult to avoid. Reinforcement learning is thus one of the most useful tools in intrusion detection to circumvent these difficulties[42], [43].

Our in-depth analysis of the complex dynamics of intrusion and defense systems consisted of a Tomcade model using a conditional entropy and constraint entropy Lagrange multipliers equation, seeking the essential knowledge that helps us develop accurate, rapid, and stable low-risk intrusion detection algorithms to predict and counteract these malicious attacks.

Reinforcement learning, which uses deep learning principles as the foundation of a mathematical architecture drawing on sequential content information drawings and incorporating special delayed reward feedback to uncover the link between actions and consequences, has two key features that make it the best method for solving anomaly-based intrusion detection as the network computer's problem [44]. It provides proactive artificial intelligence capabilities and scalable behavior for adaptive computation for the decider entities that face a continuously changing problem and strive to get the best intrusion detection and protection strategies.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

Undoubtedly, reinforcement learning has been established as the most rigorous theoretical solution to adversarial challenge cyber sequence prediction and defense game problems [45].

3.1Applications in Cybersecurity

Attacks, intrusions, and other security breaches continue to rise within technologically advanced areas such as the Internet and computers [46]. Traditional security intrusion detection systems (IDS) often use pattern-matching processes, such as signatures, and parametric or neural network models to predict whether network traffic is benign or represents a threat [47]. Such traditional IDS systems may no longer work well with the unfamiliar threats and vulnerabilities that may occur during mitigation operations like e-commerce. For these, binary decision models, such as "good" or "bad," may no longer be relevant; rather, using a more continuous measurement makes sense.

Reinforcement deep learning models use policy networks to learn from state rewards, which can be applied in theoretical and complex decision-making environments[48]. The model operates according to the Markov decision process via interacting with a reinforcement environment to select actions that optimize accumulated reward based on the deep learning process[49], [50]. A key advantage of DRL-based models is processing raw traffic data or a network flow in a whole stack of layers, measuring what is essential to characterize shape files or other data structures. They then process the raw data in the deep learning layers and use a stack as a control system for detecting and responding to intrusion/attack security incidents [50].

4. Literature Review

Deep Reinforcement Learning (DRL): Deep learning has improved several artificial intelligence techniques to translate complex perceptual tasks into training data[51]. DRL has shown promise in applying Markov Decision Processes (MDPs) with large, continuous state spaces and as a blueprinting search algorithm for achieving advanced results on real-world problems [52], [53]. However, DRL is tailored to environments that change rapidly and unpredictably, such as Go or computer games [54]. For relatively static or slowly changing environments, a computer can now search through all or very large parts of the solution space and identify the best option at any given time[55], [56]. Despite its promise, there should be more research on using DRL to identify anomalies in long-standing systems, such as network traffic.

Network Anomaly Detection: Generally speaking, the areas we have explored have different definitions, methodologies, and varying levels of efficiency[57]. These methods can be based on a logic of state and rates and construct categories with statistical analysis or use a supervised metric to compare different algorithms [58]. In either case, if these methods are deployed in network appliances, they tend to be sensitive to the network environment in which they are employed, and there is a time overhead in training and decision-making [59]. Click or tap here to enter text. They may also have a very high rate of false positives and negatives. Such obstacles in discovering anomalies may result from the current work's low efficiency and scalability in

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

identifying repositories or ground truth, which may be relatively unreliable. Achieving an effective solution requires a large-scale collection of trace data, which primarily relies on installing a networking area within a specific data type of network area or environment [57].

Cyberattacks against software based on ML have grown in frequency due to the increasing use of ML in cybersecurity[60], [61] . Cyber-attack simulation aims to train Machine Learning schemes to identify and react to real-time assaults by simulating and modeling hypothetical intrusions on a system. By following this plan, researchers may strengthen their cybersecurity measures and be better prepared to deal with cyberattacks. However, there are limitations to traditional ML-based applications, such as their reliance on historical data for training and their potential lack of generalizability[62], [63]. Recently, RL has become more popular in cybersecurity cyber-attack simulations[64], [65]. Because cyberattacks are becoming more complex and effective defenses are becoming more difficult to implement, researchers have begun to use ML techniques like RL to build security systems that are both more robust and more adaptable. The most effective means of attack mitigation, threat modification, and technique improvement may all be taught using RL algorithms. A real-life agent (RL) might learn to combat various assaults, including zero-day vulnerabilities, by regularly playing offensive and defensive games [66]. Across many scenarios, including web apps, malware research, and detecting intrusions into networks, this approach has proven helpful in discovering and containing virtual assaults[65], [67]. Cybersecurity may be drastically altered by the advent of RL-powered adaptive and automated defense systems that might gain knowledge from past mistakes and respond instantly to new cyber dangers (30). We must tackle many major challenges before properly applying RL in cybersecurity. Not having any training data is a major obstacle [68]. Cybersecurity threats are always changing, but ongoing efforts to use RL to strengthen defenses exist with an emphasis on smart grid security.

Malika et al. (2023) Anomaly-NIDS based on deep reinforcement learning is introduced in this study. method of data collection and pre-processing might be helpful in many network topologies. This reinforcement learning technique provides multiple options. The Learning mode improves the model's ability to correctly detect incoming network traffic by constantly learning and updating itself, while the detection mode optimizes processing performance. As part of the campus networking environment, the author tested the solution on 100 million Palo Alto network logs, which was effective. Three machine-learning methods were used to test the proposed DRL. Based on experimental results, the proposed method effectively achieves maximum detection accuracy, outstanding processing speed, and continuous model update efficiency [71]. However, there are no specific data for pre-processing[13].

Roger et al. (2022) This study presents a system for intrusion detection that relies on reinforcement learning and can go long periods without receiving any updates. Two approaches are incorporated into the proposal. The system employs machine learning methods as a reinforcement learning task to provide high reliability and classification accuracy over the long run. Secondly, model updates minimize computational resources and

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

human involvement using transfer learning with a sliding window technique. Experiments with an 8TB dataset and four years of actual network traffic proved that existing methods in the literature are ill-equipped to deal with the ever-changing nature of network traffic. Regarding accuracy, the method that does not update the model periodically is on par with conventional detection techniques that update their models semiannually. When updated periodically, the suggested model uses just 7 days of training data and approximately five times fewer computing resources than existing techniques while reducing accuracy variance by 6%, false positives by 8%, and false negatives by 34%[73].

Mahdi et al. (2024)
The proposed system uses sequential packet labeling to calculate an attack probability score for each fow by watching and updating each packet's estimate. The framework is evaluated using CNN-based and LSTM-based deep models on the CICIDS2017 and CSE-CIC-IDS2018 datasets. The researcher demonstrates that the suggested distributed system effectively addresses traffic idea drift via thorough evaluations and tests. Findings suggest that CNN-based models adapt to traffic concept drift, achieving above 95% detection rates with only 128 new frames. In contrast, LSTM-based models excel at sequential packet labeling in online IDSs, detecting intrusions within 15 packets[74].

Singh et al. (2024) Intelligent Intrusion Prevention System A novel approach to intrusion detection that relies on Deep Reinforcement Learning (AID-DRL) has recently been developed. An adaptable and effective intrusion detection system (IDS) that can protect against developing cyber threats is created using reinforcement learning and deep neural networks in the suggested system. In designing and constructing AID-DRL, scalability, flexibility, and integration with cybersecurity infrastructure were considered. According to the experimental data, the AID-DRL system outperformed baseline models in real-time threat detection and mitigation. Future research should focus on improving learning algorithms, making them more adversary robust, developing methods to modify policies dynamically, integrating threat intelligence, making them more scalable, improving their deployment, and finding ways to preserve privacy. These areas are designed to improve intrusion detection systems and tackle the ever-changing cybersecurity landscape[75].

Jeffrey et al. (2024) provide an approach to CPS anomaly detection that uses “unsupervised learning models” with one-class classification algorithms. This will help compensate for the extremely low amount of anomalous data included in the typically used studies. This method focuses on the differences between supervised and unsupervised learning with a restricted range of classification algorithms; nevertheless, it is not very transferable to CPS scenarios, even if it helps with some of the accuracy issues caused by imbalanced data classes [76].

Afrifa et al. (2023) begin with the premise that bad actors routinely hijack large quantities of IoT devices and turn them into botnets to accomplish their evil plans, endangering international trade. While an individual

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

hijacked IoT device would not pose much of a threat due to the limited resources available to most of these devices, a botnet consisting of hundreds or even millions could do much damage. Using Ensemble Learning, we provide a new method for detecting botnets and preventing intrusions in real-time by identifying individual nodes inside a botnet. Instead of the more typical classification job of determining if an action against a single host is harmful or benign, this novel method employs Ensemble Learning to determine whether a host is a botnet member[77].

Yazdinejad et al. (2023) provide an ensemble deep learning-based anomaly detection model for IIoT settings, which analyses time series data using the AutoEncoder (AE) architecture and Long Short-Term Memory (LTSM) to spot unusual behavior. Problems with unbalanced datasets, which impact the predictive capabilities of several ML algorithms, are prevalent in IIoT/CPS anomaly detection settings. Based on the premise that IIoT environments are dispersed and filled with diverse sensors and actuators, this study tackles the issue as a big data challenge by applying pattern recognition to time series data collected from "IIoT environment" monitoring and then determining if the activity is normal or abnormal[78].

Danso et al. (2022) To circumvent resource limitations on IoT devices, we suggest an ensemble-based intrusion detection system (IDS) installed on the IoT gateway. This system would work by passively sniffing network traffic to gather samples, which would then be utilized to train various machine learning (ML) methods to act as base learners. These models would then be fed into an ensemble learning model, improving predictive performance by combining the individual" ML algorithms with a stacking meta-classifier" to generate final anticipated results[79].

Zhao et al.(2022) present an innovative Ensemble Learning algorithm that can detect anomalies on "smart power grids". This algorithm uses "feature matching" in a system for cooperative learning to distinguish between physical faults (such as power line breaks caused by weather) and malicious actors' actions (like network-based attacks). Using an ensemble model incorporating several base classifiers, the suggested model depicts the "smart power grid" as a state machine. Anomaly behavior is then detected by processing transitions between normal states[80].

Nicholas et al(2024) This paper proposes a hybrid anomaly detection approach for CPSs, combining signature-based IT network detection, threshold-based OT network detection, and behavioral-based Ensemble Learning (EL) to enhance accuracy.
The hybrid approach is validated using multiple public research datasets. It uses a "divide-and-conquer strategy" to offload cyber threat detection to computing inexpensive signature-based and threshold-based methods, minimizing a measure of behavioral-based data for "ML model" training, resulting in higher accuracy in less time. The experiment results showed a 4-7% accuracy increase in anomaly detection across several

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

datasets, which is crucial for CPS operators because of the high financial issues and safety costs of system interruptions[81].

Vincent et al (2024) Integrated communication and information technologies with large-scale power grids have improved hyperphysical system efficiency, safety, and economy.

The smart grid is vulnerable to cyber-attacks despite its open and diverse communication environment. Data integrity attacks that overcome traditional security measures pose a significant danger to grid operations. Current detection methods for smart grids cannot adapt to dynamic and diverse features or handle non-Euclidean data sources. The author provides a new Deep-Q-Network technique using a graph convolutional network (GCN) architecture to identify data integrity breaches in cyber-physical systems. The simulation results demonstrate the framework's scalability and superior detection accuracy compared to previous benchmark methodologies[82].

Table (1) compares the selected IDSs' essential attributes in depth. However, few researchers have implemented DRL for intrusion detection. We conclude that the research and methods presented in this study have the potential to maximize detection accuracy while limiting false alarms, which will save time and money while producing trustworthy findings.

Table 1

NO	Author year	Method	Dataset	Objectives	Advantage	drawbacks
[83]	Sang et al. 2024	DRL	Simulation with Real scenario	Enhancement Realistic in CPS Simulations	efficiency improvement	dependent on similarity it's possible that they don't accurately reflect the complexity of actual cyberthreats.
[71]	Malik et al.(2023)	ANID	Palo Alto System Network Logs	Enhance speed of processing	Superior Speed processing	Restricted for certain network area
[73]	Roger et al.(2023)	RL-Based ID	8TB Real Network Traffic Data	Decrease the amount of computational Resources	High Reliability with Accuracy	used only one Dataset that is make difficult to apply to another Dataset

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

NO	Author year	Method	Dataset	Objectives	Advantage	drawbacks
[84]	Soltani et al.(2024)	multi-Agent NIDS	CICID2017&CSE-CIC-IDS2018	By using observations from several sensors, the goal is to allow a distributed IDS architecture that increases detection accuracy and addresses big data issues.	For high-throughput networks, its dispersed detection feature makes it scalable.	The implementation becomes more sophisticated with the usage of multi-agent systems
[75]	Singh et al.(2024)	DRL	NSL-KDD	Enhance accuracy ,the system penalises false positives	the system can be capable of adapting to new and emerging threats.	the method may be more complexit to implementation
[81]	jeffrey et al.(2024)	Hybrid ID	Edge-IIoTest2023 & CICIoT2023	the accuracy optimizing of ID in cyber physical systerm	Enhancement accuracy	Restricted opportunities for transfer learning
[82]	Vincent et al.(2024)	DQN & GCN		detect and prevent smart grid data integrity threats,	the accuracy detection is very high	the method may be more complexit to implementation
[77]	Afrifa et al.(2023)	ML techniques	public dataset	detect and prevent botnet attacks on linked PCS especially in the realm of IOT gadgets	The technique uses real-time AI-powered behavioural analysis to detect botnet attacks. This may be crucial for quick prevention.	Using an ensemble technique with several ML models could make implementation more complicated, necessitating additional computational resources and knowledge.
[78]	Yazdinejad et al.(2023)	Ensemble DL	GP & SWaT	Enhancement Anomaly detection in IIOT	the accuracy is very high	Complex implementation of the LSTM/AE ensemble model may necessitate substantial computer resources and deep learning knowledge.
[79]	Danso et al.(2022)	intelligent Ensemble-based IDS	CIC-IDS2017 & N-BaIoT	Enhancement accuracy detection in IOT security	the detection rate increases	Relying on Defined Datasets

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

NO	Author year	Method	Dataset	Objectives	Advantage	drawbacks
[80]	Zhao et al.(2022)	ELA		Locate and Categorise Security Vulnerabilities	Improvements in Ranking and Classification	Absence of Contextual Dataset Details

5. Methodology

This section details the method used in this research, including Deep RL algorithms. it also breaks out the basic components and how this algorithm works.

5.1. Deep Reinforcement Learning

Deep reinforcement learning methods can help improve many fields, including gaming, robotics, and natural language processing [85]. The recent study of these methods and their capabilities in solving high-dimensional data problems has made it increasingly significant to use and apply these methods to new tasks [86]. In machine learning, we have studied reward-based learning, including supervised learning, where learning occurs when training and cross-validate lists on labels from training data, as well as unsupervised learning, where learning contrasts in determining hidden patterns in the data. Reinforcement learning operates on optimizing agents that decide on sequences of actions that result in an amount of reward. It's used in domains that incorporate interacting with the environment [46].

When making decisions, an agent interacts with an environment and takes action. The environment evolves to a different state in the aftermath of each decision. After that, the agent receives a reward according to the transition of the environment or state. Future transitions generated by subsequent decisions are also based on the previous results. Based on the reward, the agent learns to associate future rewards with past decisions. This happens through the action of an agent's policy, which describes how decisions are made[46]. The learning process is caused by examining the association between rewards and the decisions made. The policy is the outcome of training towards leading to actions that maximize reward. The agent continues to cycle through experiencing decisions and observing the feasibility of the policy in yielding optimal rewards in the feedback for the process to quit[87] .

The algorithm of the decision-making framework known as reinforcement learning (RL) teaches agents how to maximize cumulative rewards through their interactions with the environment. The mathematical equation of RL is.

DRL Algorithm.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left[R_t + r^{V^{\pi}}(S_t + 1) | S_t = s\right]$$

$V^{\pi}(s)$ = value of the state (s) under the policy π.

$\mathbb{E}_{\pi}$ = an expectation that the agent follows the policy π.

R_t = reward received at time step t.

r = discount factor where ($0 \leq r \leq 1$).

$(S_t + 1)$ = state of the next time step.

The proposed method and its analysis are depicted as follows.

Our proposed AIDN is an anomaly detection engine that uses reinforcement learning. Its goal is to automatically learn new forms of attacks and network traffic behavior so it can self-update its detection model. While most intrusion detection system (IDS) research relies on hypothetical data sets, our suggested model can run in a production network setting.
That is why our technology takes processing speed and accuracy into account. We provide a method whereby data from network traffic is seen as a state variable of the environment in RL, where the RL agent is an intrusion-detecting engine. The action is the same as the result of detecting intrusions, and the reward is set according to the accuracy of the recognition result. The schematic of this system is learning and detection modes of reinforcement learning, which are necessary for the self-update function to be achieved. Here is how these two means of transportation work together:

traffic on the network data. Two, maintain the process by providing the RL agent with a fake reward via the reward function module. Revert to the first step.

Learning mode: - (i) After receiving raw data about network traffic, the RL agent processes it and then takes action based on the state variables. (ii) Rewards are computed by the reward function module and returned to the RL agent depending on the activity and the label. (iii) The RL agent uses the reward and the states to modify its policy, an intrusion detection model. Bring it all back to square one. seen in Figure (1).

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

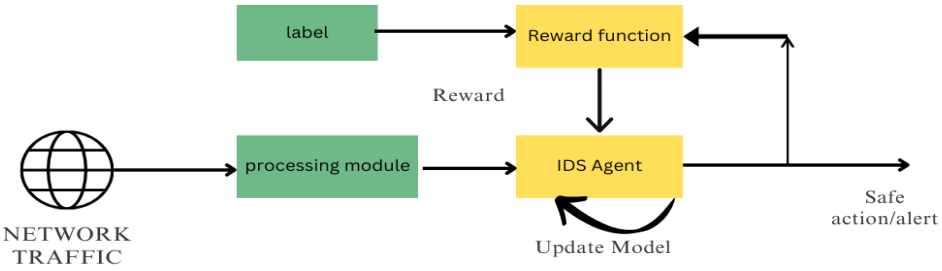
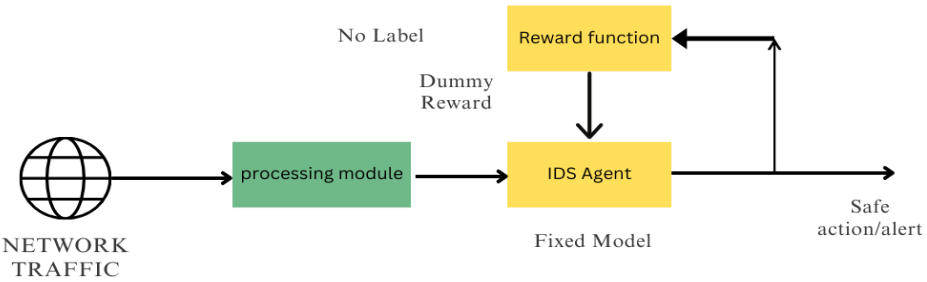


Fig (1) Learning mode

Detection mode: - (i) After receiving raw data about network traffic, the RLagent processes it and provides an action based on the state variables. (ii) the rewards module offers the RL agent a fake reward to keep the process running.(iii) return to the first step. seen in Figure (2).

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793



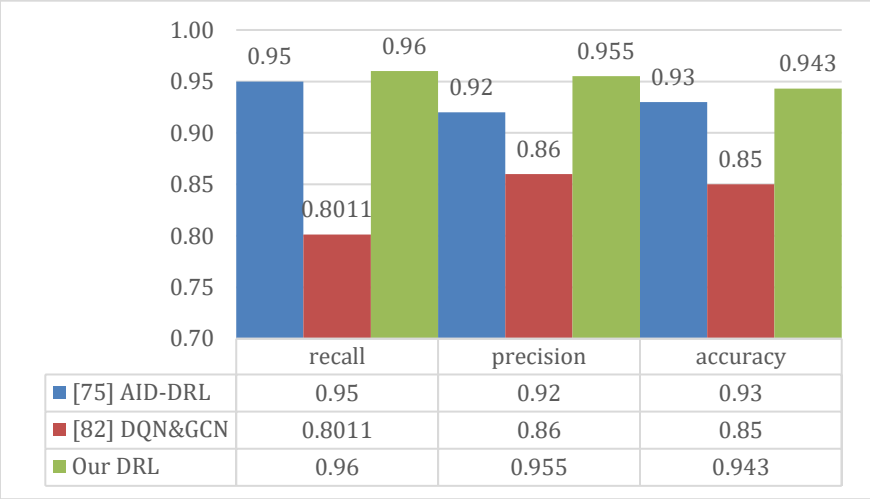
Fig(2) Detection mode

By monitoring the reward, the DRL agent may gauge how well the detection result performed while in learning mode. If the incentive is reduced, it will refresh the detection system with up-to-date information to enhance the effectiveness of intrusion detection. When operating in detection mode, the RL agent processes network data using a predefined detection model; the incentive is a fake reward that serves just to maintain the operation process. In all cases, the determining factor is whether the reward function uses the label to determine the actual payout. We set a switch flag to allow the system to transition between the two modes with flexibility. Thanks to this configuration, the system may assess and upgrade the detection model at any moment. As an intrusion detection result, the DRL agent will take a specific action; our system may also extend to avoid intrusions.

In Table (2) , our suggested DRL model achieves better results than competing baseline methods in attack classification, with a recall of 96.0% and an accuracy of 94.3%. Our study's three models, AID-DRL, DQN-GCN, and OUR, proposed that all of these achieve better accuracy thanks to their reduced false-positive ratio. Still, their recall rates (ranging from 80.1% to 96.0%) could be due to the significant amount of false-negative predictions they make. Mistaking unusual data transmissions for regular ones is what the false-negative symbolizes. Hence, it raises the possibility of network infiltration and lowers system dependability. Our method attains a balance of almost 94% across the assessment criteria of recall, precision, and accuracy, as shown by the experimental results.

Table 2

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793



	recall	precision	accuracy
[75] AID-DRL	0.95	0.92	0.93
[82] DQN&GCN	0.8011	0.86	0.85
Our DRL	0.96	0.955	0.943

6. Conclusion and future work.

Anomaly detection and behavior modeling-based detection methods are known for their capabilities to identify unknown or novel intrusions. These methods first discover the "normal/usual" processes or system behaviors. They establish models of those discovered processes or behaviors to represent how the system functions typically or behaves. After that, those methods detect unknown or novel intrusions by examining the monitored system activities.

To protect the confidentiality, availability, and integrity of data and communications, network intrusion detection systems (NIDSs) monitor network traffic for signs of malicious activity or cyberattacks. We provide an anomalous network intrusion detection system (ANIDS) that uses deep RL to identify DRL-related malicious network behavior. We provide an approach that incorporates data collecting and preprocessing stages and may be used to diverse network setups. In this reinforcement learning approach, you are given several options. In contrast to the detection mode, which prioritizes speed, the learning mode is designed to train and update the model to achieve the greatest possible accuracy for continuous streams of data arriving from a network. Then, to prove it works even better, we used our method and pitted our suggested DRL against two ML models. The proposed model outperforms competing techniques regarding processing speed, detection accuracy, and long-term model effectiveness. Future work in this field will be very complex. However, although there is great potential for DRL integration in cybersecurity, some constraints and problems must be overcome to make these models more successful.
The evolution of more complex simulation environments capable of responding to real-time changes in attack

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

strategies and network configurations is a significant obstacle due to the ever-changing nature of cyber threats. In the future, research should focus on finding ways to reduce the resource consumption of DRL training. Possible solutions might include using shared parameters or compressing models. Researching effective transfer learning procedures is also necessary to improve the reusability of DRL models in both simulated and real-world settings. Additionally, future research should concentrate on adversarial safeguards, interpretable DRL models, and continuous learning architectures to guarantee that automated cybersecurity decision-making complies with ethical and legal standards; validation and assurance frameworks should be developed. These projects and the complex problems at the junction of DRL and cybersecurity can only be solved via the joint work of researchers in machine learning and cybersecurity.

Bibliogrphy.

[1] M. Abdel-Rahman and others, "Advanced cybersecurity measures in IT service operations and their crucial role in safeguarding enterprise data in a connected world," *Eigenpub Review of Science and Technology*, vol. 7, no. 1, pp. 138–158, 2023.

[2] A. Heidari and M. A. Jabraeil Jamali, "Internet of Things intrusion detection systems: a comprehensive review and future directions," *Cluster Comput*, vol. 26, no. 6, pp. 3753–3780, 2023.

[3] M. Abdulhussein, *The Impact of Artificial Intelligence and Machine Learning on Organizations Cybersecurity*. Liberty University, 2024.

[4] A. Thakkar and R. Lohiya, "A review on challenges and future research directions for machine learning-based intrusion detection system," *Archives of Computational Methods in Engineering*, vol. 30, no. 7, pp. 4245–4269, 2023.

[5] R. I. Mukhamediev *et al.*, "Review of artificial intelligence and machine learning technologies: classification, restrictions, opportunities and challenges," *Mathematics*, vol. 10, no. 15, p. 2552, 2022.

[6] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, 2021.

[7] G. Kocher and G. Kumar, "Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges," *Soft comput*, vol. 25, no. 15, pp. 9731–9763, 2021.

[8] E. Schiller, A. Aidoo, J. Fuhrer, J. Stahl, M. Ziörjen, and B. Stiller, "Landscape of IoT security," *Comput Sci Rev*, vol. 44, p. 100467, 2022.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[9] Y. Bin Zikria, R. Ali, M. K. Afzal, and S. W. Kim, "Next-generation internet of things (iot): Opportunities, challenges, and solutions," *Sensors*, vol. 21, no. 4, p. 1174, 2021.

[10] C. P. Kaliappan, K. Palaniappan, D. Ananthavadivel, and U. Subramanian, "Advancing IoT security: a comprehensive AI-based trust framework for intrusion detection," *Peer Peer Netw Appl*, pp. 1–21, 2024.

[11] V. Demertzi, S. Demertzis, and K. Demertzis, "An overview of cyber threats, attacks and countermeasures on the primary domains of smart cities," *Applied Sciences*, vol. 13, no. 2, p. 790, 2023.

[12] I. A. I. Ahmad, A. C. Anyanwu, S. Onwusinkwue, S. O. Dawodu, O. V. Akagha, and E. Ejairu, "Cybersecurity challenges in smart cities: a case review of African metropolises," *Computer Science & IT Research Journal*, vol. 5, no. 2, pp. 254–269, 2024.

[13] M. Sarhan, S. Layeghy, M. Gallagher, and M. Portmann, "From Zero-Shot Machine Learning to Zero-Day Attack Detection. arXiv 2021," *arXiv preprint arXiv:2109.14868*.

[14] D. Nair and N. Mhavan, "Augmenting Cybersecurity: A Survey of Intrusion Detection Systems in Combating Zero-day Vulnerabilities," in *Smart Analytics, Artificial Intelligence and Sustainable Performance Management in a Global Digitalised Economy*, Emerald Publishing Limited, 2023, pp. 129–153.

[15] S. Applebaum, T. Gaber, and A. Ahmed, "Signature-based and machine-learning-based web application firewalls: a short survey," *Procedia Comput Sci*, vol. 189, pp. 359–367, 2021.

[16] M. Alanazi, A. Mahmood, and M. J. M. Chowdhury, "SCADA vulnerabilities and attacks: A review of the state-of-the-art and open issues," *Comput Secur*, vol. 125, p. 103028, 2023.

[17] Y. Guo, "A review of Machine Learning-based zero-day attack detection: Challenges and future directions," *Comput Commun*, vol. 198, pp. 175–185, 2023.

[18] B. Gao *et al.*, "Enhancing anomaly detection accuracy and interpretability in low-quality and class imbalanced data: A comprehensive approach," *Appl Energy*, vol. 353, p. 122157, 2024.

[19] R. Liu, J. Shi, X. Chen, and C. Lu, "Network anomaly detection and security defense technology based on machine learning: A review," *Computers and Electrical Engineering*, vol. 119, p. 109581, 2024.

[20] L. Bergman and Y. Hoshen, "Classification-based anomaly detection for general data," *arXiv preprint arXiv:2005.02359*, 2020.

[21] R. Jiao *et al.*, "Learning with limited annotations: a survey on deep semi-supervised learning for medical image segmentation," *Comput Biol Med*, p. 107840, 2023.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[22] J. Zipfel, F. Verworner, M. Fischer, U. Wieland, M. Kraus, and P. Zschech, "Anomaly detection for industrial quality assurance: A comparative evaluation of unsupervised deep learning models," *Comput Ind Eng*, vol. 177, p. 109045, 2023.

[23] M. Z. Zaheer, A. Mahmood, M. H. Khan, M. Segu, F. Yu, and S.-I. Lee, "Generative cooperative learning for unsupervised video anomaly detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 14744–14754.

[24] A. Azab, M. Khasawneh, S. Alrabaee, K.-K. R. Choo, and M. Sarsour, "Network traffic classification: Techniques, datasets, and challenges," *Digital Communications and Networks*, vol. 10, no. 3, pp. 676–692, 2024.

[25] G. ALMahadin *et al.*, "VANET network traffic anomaly detection using GRU-based deep learning model," *IEEE Transactions on Consumer Electronics*, 2023.

[26] D. Javaheri, S. Gorgin, J.-A. Lee, and M. Masdari, "Fuzzy logic-based DDoS attacks and network traffic anomaly detection methods: Classification, overview, and future perspectives," *Inf Sci (N Y)*, vol. 626, pp. 315–338, 2023.

[27] R. Al-amri, R. K. Murugesan, M. Man, A. F. Abdulateef, M. A. Al-Sharafi, and A. A. Alkahtani, "A review of machine learning and deep learning techniques for anomaly detection in IoT data," *Applied Sciences*, vol. 11, no. 12, p. 5320, 2021.

[28] M. H. Thwaini, "Anomaly Detection in Network Traffic using Machine Learning for Early Threat Detection," *Data and Metadata*, vol. 1, p. 72, 2022.

[29] L. I. Khalaf, B. Alhamadani, O. A. Ismael, A. A. Radhi, S. R. Ahmed, and S. Algburi, "Deep Learning-Based Anomaly Detection in Network Traffic for Cyber Threat Identification," in *Proceedings of the Cognitive Models and Artificial Intelligence Conference*, 2024, pp. 303–309.

[30] M. Abbasi, A. Shahraki, and A. Taherkordi, "Deep learning for network traffic monitoring and analysis (NTMA): A survey," *Comput Commun*, vol. 170, pp. 19–41, 2021.

[31] Z. Yang *et al.*, "A systematic literature review of methods and datasets for anomaly-based network intrusion detection," *Comput Secur*, vol. 116, p. 102675, 2022.

[32] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine learning for anomaly detection: A systematic review," *Ieee Access*, vol. 9, pp. 78658–78700, 2021.

[33] X. Ma *et al.*, "A comprehensive survey on graph anomaly detection with deep learning," *IEEE Trans Knowl Data Eng*, vol. 35, no. 12, pp. 12012–12038, 2021.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[34] D. K. Reddy, H. S. Behera, J. Nayak, P. Vijayakumar, B. Naik, and P. K. Singh, "Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 7, p. e4121, 2021.

[35] D. Ageyev, T. Radivilova, O. Mulesa, O. Bondarenko, and O. Mohammed, "Traffic monitoring and abnormality detection methods for decentralized distributed networks," in *Information security technologies in the decentralized distributed networks*, Springer, 2022, pp. 287–305.

[36] A. Thakkar and R. Lohiya, "A review on machine learning and deep learning perspectives of IDS for IoT: recent updates, security issues, and challenges," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 3211–3243, 2021.

[37] Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4150, 2021.

[38] X. Wang *et al.*, "Deep reinforcement learning: A survey," *IEEE Trans Neural Netw Learn Syst*, vol. 35, no. 4, pp. 5064–5078, 2022.

[39] V. G. da Silva Ruffo *et al.*, "Anomaly and intrusion detection using deep learning for software-defined networks: A survey," *Expert Syst Appl*, p. 124982, 2024.

[40] G. Apruzzese *et al.*, "The role of machine learning in cybersecurity," *Digital Threats: Research and Practice*, vol. 4, no. 1, pp. 1–38, 2023.

[41] M. Lydia, G. E. Prem Kumar, and A. I. Selvakumar, "Securing the cyber-physical system: A review," *Cyber-Physical Systems*, vol. 9, no. 3, pp. 193–223, 2023.

[42] S. Mohamed and R. Ejbali, "Deep SARSA-based reinforcement learning approach for anomaly network intrusion detection system," *Int J Inf Secur*, vol. 22, no. 1, pp. 235–247, 2023.

[43] S. Tharewal, M. W. Ashfaque, S. S. Banu, P. Uma, S. M. Hassen, and M. Shabaz, "Intrusion detection system for industrial Internet of Things based on deep reinforcement learning," *Wirel Commun Mob Comput*, vol. 2022, no. 1, p. 9023719, 2022.

[44] B. Sharma, L. Sharma, C. Lal, and S. Roy, "Anomaly based network intrusion detection for IoT attacks using deep learning technique," *Computers and Electrical Engineering*, vol. 107, p. 108626, 2023.

[45] A. A. Hammad, S. R. Ahmed, M. K. Abdul-Hussein, M. R. Ahmed, D. A. Majeed, and S. Algburi, "Deep Reinforcement Learning for Adaptive Cyber Defense in Network Security," in *Proceedings of the Cognitive Models and Artificial Intelligence Conference*, 2024, pp. 292–297.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[46] A. Malekloo, E. Ozer, M. AlHamaydeh, and M. Girolami, "Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights," *Struct Health Monit*, vol. 21, no. 4, pp. 1906–1955, 2022.

[47] S. Thudumu, P. Branch, J. Jin, and J. Singh, "A comprehensive survey of anomaly detection techniques for high dimensional big data," *J Big Data*, vol. 7, pp. 1–30, 2020.

[48] M. Naeem *et al.*, "Trends and future perspective challenges in big data," in *Advances in Intelligent Data Analysis and Applications: Proceeding of the Sixth Euro-China Conference on Intelligent Data Analysis and Applications, 15–18 October 2019, Arad, Romania*, 2022, pp. 309–325.

[49] I. Lee and B. Perret, "Preparing high school teachers to integrate AI methods into STEM classrooms," in *Proceedings of the AAAI conference on artificial intelligence*, 2022, pp. 12783–12791.

[50] G. Aguiar, B. Krawczyk, and A. Cano, "A survey on learning from imbalanced data streams: taxonomy, challenges, empirical study, and reproducible experimental framework," *Mach Learn*, vol. 113, no. 7, pp. 4165–4243, 2024.

[51] S. Latif, H. Cuayáhuitl, F. Pervez, F. Shamshad, H. S. Ali, and E. Cambria, "A survey on deep reinforcement learning for audio-based applications," *Artif Intell Rev*, vol. 56, no. 3, pp. 2193–2240, 2023.

[52] X. Yang, E. Howley, and M. Schukat, "ADT: Time series anomaly detection for cyber-physical systems via deep reinforcement learning," *Comput Secur*, vol. 141, p. 103825, 2024.

[53] D. Han, "HARNESSING DEEP REINFORCEMENT LEARNING: STUDIES IN ROBOTIC MANIPULATION, ENHANCED SEMANTIC SEGMENTATION, AND SECURING IMAGE CLASSIFIERS," 2024.

[54] S. Zhang, Y. Li, F. Ye, X. Geng, Z. Zhou, and T. Shi, "A hybrid human-in-the-loop deep reinforcement learning method for UAV motion planning for long trajectories with unpredictable obstacles," *Drones*, vol. 7, no. 5, p. 311, 2023.

[55] M. Świechowski, K. Godlewski, B. Sawicki, and J. Mańdziuk, "Monte Carlo tree search: A review of recent modifications and applications," *Artif Intell Rev*, vol. 56, no. 3, pp. 2497–2562, 2023.

[56] K. Rajwar, K. Deep, and S. Das, "An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges," *Artif Intell Rev*, vol. 56, no. 11, pp. 13187–13257, 2023.

[57] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert Syst Appl*, vol. 193, p. 116429, 2022.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[58] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Comput Biol Med*, vol. 136, p. 104672, 2021.

[59] R. Ahmad, R. Wazirali, and T. Abu-Ain, "Machine learning for wireless sensor networks security: An overview of challenges and issues," *Sensors*, vol. 22, no. 13, p. 4730, 2022.

[60] R. Sen, G. Heim, and Q. Zhu, "Artificial intelligence and machine learning in cybersecurity: Applications, challenges, and opportunities for mis academics," *Communications of the Association for Information Systems*, vol. 51, no. 1, p. 28, 2022.

[61] A. Pinto, L.-C. Herrera, Y. Donoso, and J. A. Gutierrez, "Survey on intrusion detection systems based on machine learning techniques for the protection of critical infrastructure," *Sensors*, vol. 23, no. 5, p. 2415, 2023.

[62] A. Piplai, M. Anoruo, K. Fasaye, A. Joshi, T. Finin, and A. Ridley, "Knowledge guided two-player reinforcement learning for cyber attacks and defenses," in *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2022, pp. 1342–1349.

[63] V. Duddu, "A survey of adversarial machine learning in cyber warfare," *Def Sci J*, vol. 68, no. 4, p. 356, 2018.

[64] M. A. R. Al Amin, S. Shetty, and C. Kamhoua, "Cyber deception metrics for interconnected complex systems," in *2022 Winter Simulation Conference (WSC)*, 2022, pp. 473–483.

[65] H. Rathore, S. K. Sahay, P. Nikam, and M. Sewak, "Robust android malware detection system against adversarial attacks using q-learning," *Information Systems Frontiers*, vol. 23, pp. 867–882, 2021.

[66] Y. Huang, L. Huang, and Q. Zhu, "Reinforcement learning for feedback-enabled cyber resilience," *Annu Rev Control*, vol. 53, pp. 273–295, 2022.

[67] K. Sethi, Y. V. Madhav, R. Kumar, and P. Bera, "Attention based multi-agent intrusion detection systems using reinforcement learning," *Journal of Information Security and Applications*, vol. 61, p. 102923, 2021.

[68] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, "Cybersecurity threats and their mitigation approaches using Machine Learning—A Review," *Journal of Cybersecurity and Privacy*, vol. 2, no. 3, pp. 527–555, 2022.

[69] M. Ibrahim and R. Elhafiz, "Security analysis of cyber-physical systems using reinforcement learning," *Sensors*, vol. 23, no. 3, p. 1634, 2023.

[70] A. Dutta, S. Chatterjee, A. Bhattacharya, and M. Halappanavar, "Deep reinforcement learning for cyber system defense under dynamic adversarial uncertainties," *arXiv preprint arXiv:2302.01595*, 2023.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[71] M. Malik and K. S. Saini, "Network Intrusion Detection System using Reinforcement learning," in *2023 4th International Conference for Emerging Technology (INCET)*, 2023, pp. 1–4.

[72] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data," *Front Energy Res*, vol. 9, p. 652801, 2021.

[73] R. R. dos Santos, E. K. Viegas, A. O. Santin, and V. V Cogo, "Reinforcement learning for intrusion detection: More model longness and fewer updates," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 2040–2055, 2022.

[74] M. Soltani, K. Khajavi, M. Jafari Siavoshani, and A. H. Jahangir, "A multi-agent adaptive deep learning framework for online intrusion detection," *Cybersecurity*, vol. 7, no. 1, p. 9, 2024.

[75] D. N. Singh, S. Jaiswar, P. Jha, V. K. Tiwari, and P. K. Saket, "Adaptive Intrusion Detection Using Deep Reinforcement Learning: A Novel Approach," *International Journal of all Research Educvation & Scientific Methods*, vol. 12, no. 05.

[76] N. Jeffrey, Q. Tan, and J. R. Villar, "A hybrid methodology for anomaly detection in Cyber–Physical Systems," *Neurocomputing*, vol. 568, p. 127068, 2024.

[77] S. Afrifa, V. Varadarajan, P. Appiahene, T. Zhang, and E. A. Domfeh, "Ensemble machine learning techniques for accurate and efficient detection of botnet attacks in connected computers," *Eng*, vol. 4, no. 1, pp. 650–664, 2023.

[78] A. Yazdinejad, M. Kazemi, R. M. Parizi, A. Dehghantanha, and H. Karimipour, "An ensemble deep learning model for cyber threat hunting in industrial internet of things," *Digital Communications and Networks*, vol. 9, no. 1, pp. 101–110, 2023.

[79] P. K. Danso, E. C. P. Neto, S. Dadkhah, A. Zohourian, H. Molyneaux, and A. A. Ghorbani, "Ensemble-based intrusion detection for internet of things devices," in *2022 IEEE 19th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, 2022, pp. 34–39.

[80] H. Zhao, C. Li, X. Yin, X. Li, R. Zhou, and R. Fu, "Ensemble Learning-Enabled Security Anomaly Identification for IoT Cyber–Physical Power Systems," *Electronics (Basel)*, vol. 11, no. 23, p. 4043, 2022.

[81] N. Jeffrey, Q. Tan, and J. R. Villar, "Using Ensemble Learning for Anomaly Detection in Cyber–Physical Systems," *Electronics (Basel)*, vol. 13, no. 7, p. 1391, 2024.

[82] E. Vincent, M. Korki, M. Seyedmahmoudian, A. Stojcevski, and S. Mekhilef, "Reinforcement learning-empowered graph convolutional network framework for data integrity attack detection in cyber-physical systems," *CSEE Journal of Power and Energy Systems*, 2024.

Vol. 15 No. 2 (2024): 309-332
https://doi.org/10.61841/turcomat.v15i2.14793

[83] S. H. Oh, J. Kim, J. H. Nah, and J. Park, "Employing Deep Reinforcement Learning to Cyber-Attack Simulation for Enhancing Cybersecurity," *Electronics (Basel)*, vol. 13, no. 3, p. 555, 2024.

[84] M. Soltani, K. Khajavi, M. Jafari Siavoshani, and A. H. Jahangir, "A multi-agent adaptive deep learning framework for online intrusion detection," *Cybersecurity*, vol. 7, no. 1, p. 9, 2024.

[85] V. Uc-Cetina, N. Navarro-Guerrero, A. Martin-Gonzalez, C. Weber, and S. Wermter, "Survey on reinforcement learning for language processing," *Artif Intell Rev*, vol. 56, no. 2, pp. 1543–1575, 2023.

[86] L. Zhang and L. Zhang, "Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities," *IEEE Geosci Remote Sens Mag*, vol. 10, no. 2, pp. 270–294, 2022.

[87] W. Jia, M. Sun, J. Lian, and S. Hou, "Feature dimensionality reduction: a review," *Complex & Intelligent Systems*, vol. 8, no. 3, pp. 2663–2693, 2022.