Computing&AI Connect
ISSN: 3104-4719
2024, Article ID. 2024.0008, Cite as: https://doi.org/10.69709/CAIC.2024.100109

SCIFINITI

Review Article

OPEN ACCESS

Distributed Reinforcement Learning for IoT Security in Heterogeneous and Distributed Networks

Senthil Kumar Jagatheesaperumal✉,1 Mohamed Rahouti✉,1,2 Mohammed Aledhari✉,3 Abdelatif Hafid✉,4 Diogo Oliveira✉,5 Hamza Drid✉,6 Ruhul Amin✉,2

[1] Department of Electronics & Communication Engineering, Mepco Schlenk Engineering College, Sivakasi 626005, India
[2] Department of Computer and Information Science, Fordham University, Bronx, NY 10023, USA
[3] Department of Information Science, University of North Texas, Denton, TX 76203, USA
[4] ESISA Analytica, Higher School of Engineering in Applied Sciences, Fez 30050, Morocco
[5] Applied Computer Science, Point Park University, Pittsburgh, PA 15222, USA
[6] Department of Computer Science, University of Batna 2, Batna 05078, Algeria

Article History

Submitted: August 01, 2024 Accepted: December 04, 2024 Published: December 31, 2024

Abstract

The explosive growth of the Internet of Things (IoT) has significantly increased networked devices within distributed and heterogeneous networks. Due to these networks' inherent vulnerabilities and diversity, the proliferation of IoT devices presents substantial security challenges. Traditional security solutions face challenges in keeping up with the constantly changing threats in dynamic situations. This article reviews the application of distributed Reinforcement Learning approaches to enhance IoT security in dispersed and heterogeneous networks. This paper provides a comprehensive overview of the fundamental theories reinforcing IoT security. It also explores the basis of Distributed Reinforcement Learning and discuss its benefits and drawbacks for IoT security. Then, the focus is given on how Distributed Reinforcement Learning might address these issues and offer details on the design factors to consider when implementing Distributed Reinforcement Learning-based solutions into practice. The paper outlines case studies and experiments that show how Distributed Reinforcement Learning may enhance IoT security. It also addresses performance analysis and evaluation measures to compare Distributed Reinforcement Learning-based approaches with conventional security methods. Finally, the paper highlights the possible uses of Distributed Reinforcement Learning in IoT security and suggest future directions, emerging trends, and unresolved challenges.

Keywords:

IoT security; Distributed Learning; Reinforcement Learning; artificial intelligence; distributed networks; heterogeneous networks

1. Introduction

The data stream from the Internet of Things (IoT) devices is vulnerable to cyber risks, a significant concern for end users. This susceptibility could also lead to the potential misguidance of both users and ML models with erroneous information during their training and learning phases. Consequently, policymakers and industries have recently started recognizing that expanding interconnected devices and their susceptibility to cyber threats introduces substantial risks, encompassing malicious attacks and the potential for distorted inferences [1]. This recognition has prompted the need to identify and comprehend these risks to develop effective security solutions. Furthermore, the economic repercussions of IoT devices and the associated security vulnerabilities are escalating in parallel with the integration of artificial intelligence (AI) into human-computer interaction (HCI) domains, including sectors like banking and insurance [2].

*** Corresponding Author:**
Mohamed Rahouti, Department of Computer and Information Science, Fordham University, Bronx, NY 10023, USA, mra-houti@fordham.edu

The rapid expansion of IoT has profoundly impacted how people interact with technology. It has facilitated the connection of numerous devices, allowing for unprecedented data sharing and automation. The Internet of Things has a transformative impact on many industries. This impact can be seen in smart homes, wearables, industrial systems, and medical equipment [3]. However, this remarkable growth also introduces many security challenges that must be addressed to safeguard sensitive information, uphold privacy, and prevent malicious activities.

Securing networked systems in the face of IoT device proliferation within dispersed and heterogeneous networks poses a significant challenge. The complexity of this task is amplified by the diversity of capabilities, communication protocols, and operating systems exhibited by IoT devices [4–6].

This diversity expands the attack surface, leaving vulnerabilities that adversaries can exploit. Furthermore, the distributed nature of IoT networks, spanning various contexts and locales, presents formidable obstacles to effectively implementing centralized security measures [7, 8]. The scattered deployment of IoT devices makes enforcing uniform security protocols challenging, necessitating exploring alternative approaches to enhance the protection of these interconnected systems.

Conventional security solutions, designed primarily for traditional IT networks, often do not adequately address the unique security concerns posed by IoT installations. The static nature of conventional security methods proves insufficient in dynamic IoT ecosystems, where devices continuously join and depart from the network [9]. Moreover, the ever-evolving threat landscape demands security systems that are flexible and adaptable, capable of responding promptly to emerging dangers. To effectively protect IoT environments, innovative security approaches that can accommodate the inherent dynamism and evolving nature of IoT systems are imperative [10]. Such adaptive security measures are essential to ensure the resilience and robustness of IoT deployments in the face of emerging threats.

Hence, there is an urgent imperative to enhance IoT security in heterogeneous and dispersed networks, enabling businesses and consumers to harness the potential of IoT technology while fortifying their defenses against potential attacks. This review paper investigates the possibility of using Distributed Reinforcement Learning techniques as a viable approach to address these security challenges and protect IoT devices against the evolving landscape of cyber threats [11]. By leveraging the power of Distributed Reinforcement Learning, the aim is to explore novel avenues for bolstering the security of IoT deployments and establishing robust protection mechanisms to ensure the integrity and resilience of IoT systems in the face of emerging risks.

1.1. Motivation

The proliferation of IoT devices has transformed connectivity, enabling automation and data-driven decision-making across critical sectors such as healthcare, energy, and transportation. However, this remarkable growth comes with substantial challenges. IoT networks are inherently distributed and heterogeneous, characterized by diverse device capabilities, communication protocols, and deployment contexts. This diversity creates significant security vulnerabilities, expanding the attack surface and exposing IoT systems to threats like unauthorized access, data breaches, and Denial-of-Service (DoS) attacks.

Additionally, integrating AI into IoT systems introduces new risks, as these technologies rely heavily on data integrity for accurate decision-making. Adversarial attacks or data manipulation can lead to cascading failures, jeopardizing critical infrastructure and services.

Traditional centralized security solutions are insufficient for such scenarios due to their inability to adapt to the dynamic and decentralized nature of IoT environments. This urgent need for scalable, adaptive, and resource-efficient security mechanisms motivates the exploration of Distributed Reinforcement Learning as a promising solution.

1.2. Problem Statement

IoT networks face a unique combination of challenges that exacerbate their vulnerability to cyber threats:

- Heterogeneity: IoT networks encompass devices with diverse hardware, communication protocols, and software configurations, complicating the enforcement of uniform security measures.
- Dynamic topologies: IoT devices frequently join and leave networks, leading to constantly evolving configurations that make traditional, static security mechanisms obsolete.
- Resource constraints: Many IoT devices operate with limited processing power, memory, and energy, rendering them incapable of supporting complex security protocols.
- Evolving threat landscape: IoT systems are increasingly targeted by sophisticated cyberattacks, including adversarial manipulations and data poisoning, requiring security frameworks to adapt in real-time.

Addressing these challenges necessitates a paradigm shift from static, rule-based solutions to adaptive, decentralized approaches. This study seeks to leverage Distributed Reinforcement Learning to develop a robust framework capable of dynamically securing IoT networks. Distributed Reinforcement Learning enables decentralized decision-making and adaptive learning, making it an ideal candidate for addressing the complexities of IoT security.

1.3. Contributions

This paper makes the following key contributions:

- Provides a detailed survey of existing Distributed Reinforcement Learning approaches for IoT security, highlighting their strengths, limitations, and applicability in heterogeneous and distributed networks.
- Identifies key challenges in IoT security, including scalability, adaptability, resource constraints, and resistance to adversarial attacks, providing a foundation for future research.
- Offers a structured categorization of DRL techniques and their applications in IoT, aiding researchers in navigating the existing literature.
- Presents a DRL-based framework specifically designed to enhance IoT security, focusing on scalability, adaptability, and energy efficiency.
- Discusses essential evaluation metrics such as detection accuracy, response time, and robustness and applies these to assess the proposed framework.
- Future Research Directions: Highlights open research gaps and emerging trends in the intersection of DRL and IoT security, providing a roadmap for innovation in the field.

1.4. Related Papers

Table 1 compares the attached survey paper [12] with recent literature in IoT security. While existing works have provided significant insights into applying machine learning and deep learning for IoT security, [12] distinguishes itself by focusing on emerging technologies such as generative AI and large language models (LLMs). This forward-looking perspective fills an important gap in the literature by addressing the potential of these technologies to transform IoT security.

Table 1: Comparison of our paper with state-of-the-art surveys [12–21].

Ref.	Focus Area	Strengths	Limitations
[12]	ML techniques for IoT security with a focus on generative AI and LLMs	Future-oriented vision with the inclusion of emerging technologies such as generative AI; discusses LLMs' role in IoT security	Lacks detailed empirical evaluation of proposed approaches
[13]	IoT security using ML; trends and challenges	Focus on IoT-specific challenges and ML solutions; comprehensive discussion of industrial applications	Limited coverage of generative AI and its potential integration with ML in IoT
[14]	Detection of DDoS attacks in IoT networks using DL and feature fusion	Detailed review of DDoS-specific detection methods; highlights the role of feature fusion in improving detection accuracy	Narrow focus on DDoS; limited generalization to other IoT threats
[15]	Applications of distributed ML in IoT	Comprehensive survey of distributed ML techniques for IoT; includes scalability challenges	Lacks specific focus on RL or security-specific applications
[16]	RL for IoT security	Comprehensive survey of RL methods applied to IoT security; identifies gaps and challenges in RL for IoT	Focuses primarily on RL without exploring hybrid ML approaches or federated learning
[17]	DRL in IoT	Detailed discussion of DRL techniques; includes use cases and performance evaluations for IoT scenarios	Overlooks generative AI and recent advancements in hybrid methodologies
[18]	ML and DL methods for IoT security	Broad coverage of ML and DL techniques for IoT security; detailed categorization of approaches	Lacks focus on RL and emerging AI technologies

Table 1: *Cont.*

Ref.	Focus Area	Strengths	Limitations
[19]	RL and DRL in IoT	Covers RL and DRL applications in IoT, focusing on wireless IoT systems	Limited discussion of IoT security challenges; focuses on wireless communication
[20]	Distributed ML in wireless communication networks	Highlights techniques and architectures for distributed ML in wireless IoT networks	Does not directly address IoT security concerns or RL approaches
[21]	ML for IoT security	Systematic literature review of ML methods for IoT security; includes taxonomy and evaluation metrics	Does not consider generative AI or advanced ML approaches like DRL

Papers such as [13,14] provide valuable discussions on specific aspects of IoT security, such as IoT challenges and DoS detection. However, they lack the broader vision offered by [12] in terms of incorporating next-generation AI approaches. Similarly, refs. [16,19] focus exclusively on reinforcement learning methods, which, while impactful, do not cover integrating hybrid AI techniques or recent advancements in generative models.

On the other hand, refs. [17,20] highlight the potential of distributed and deep reinforcement learning in IoT networks but fall short in exploring their implications for emerging AI technologies like LLMs. Compared to these works, ref. [12] provides a more comprehensive future vision, making it a more insightful and valuable contribution to the field.

1.5. Paper Organization

The rest of this article is organized as follows: Section 2 introduces the background details of IoT security and Distributed Reinforcement Learning. Section 3 explores the usage of Distributed Reinforcement Learning for IoT security. Section 4 presents the most common datasets and some use cases of Distributed Reinforcement Learning in IoT security. Section 5 highlights the evaluation metrics and performance analysis of using Distributed Reinforcement Learning for IoT security applications. Section 6 discusses the future research directions of Distributed Reinforcement Learning in IoT security. Section 7 concludes the article with a summary of key findings.

2. IoT Security and Distributed Reinforcement Learning Fundamentals

IoT has transformed the interaction with technology by bridging the physical and digital realms. Still, it also brings significant security challenges, such as unauthorized access, data breaches, and physical threats, implementing strong IoT security measures essential for protecting privacy, data, and infrastructure. The Fundamentals of IoT security and DFL are discussed as below.

2.1. Fundamentals of IoT Security

IoT has revolutionized how people interact with technology, enabling a seamless connection between the physical and digital worlds. However, with this increased connectivity comes the need for robust security measures. IoT devices are vulnerable to various threats, including unauthorized access, data breaches, and even physical harm. Therefore, understanding and implementing the fundamentals of IoT security is crucial to safeguarding our privacy, data, and infrastructure [22–26].

2.1.1. Authentication and Authorization

Authentication and authorization are essential components of IoT security. Authentication verifies the identity of devices and users before granting access to the network [27–30]. Strong authentication mechanisms, such as two-factor authentication, biometrics, or cryptographic protocols, should be implemented to prevent unauthorized access. Authorization ensures that authenticated devices and users are granted appropriate privileges and access rights [31–33]. Access control policies, role-based access control, and secure communication protocols are all critical for effective authorization in IoT systems.

2.1.2. Secure Communication

Securing communication channels is crucial in IoT deployments to protect data integrity and confidentiality.

Encryption algorithms, such as the Advanced Encryption Standard (AES), should be utilized to secure data transmission between IoT devices and the network, ensuring confidentiality and protection against unauthorized access. Secure communication protocols like SSL/TLS provide a secure channel for data transfer, preventing eavesdropping and man-in-the-middle attacks. Additionally, virtual private networks (VPNs) can add an extra layer of security by creating encrypted tunnels for communication between IoT devices and the cloud [34].

2.1.3. Device and Firmware Security

The security of the IoT devices themselves is of utmost importance. Manufacturers should implement secure development practices, including secure coding, rigorous testing, and vulnerability assessments [35]. IoT devices should have robust security controls, such as secure boot mechanisms, tamper-proof hardware, and firmware integrity checks. Regular security updates and patches should be provided to address vulnerabilities that may arise over time. Moreover, devices should have mechanisms to reset or revoke compromised credentials and certificates.

2.1.4. Data Protection and Privacy

The Internet of Things (IoT) generates vast volumes of sensitive data, making its protection a critical priority [36]. Data should be encrypted both during transmission and at rest, ensuring that even if intercepted, it remains secure. Data access should only be granted to authorized users, and data storage should adhere to industry best practices and regulations, such as the General Data Protection Regulation (GDPR). To minimize privacy risks associated with individual data points, it is possible to utilize anonymization and aggregation techniques [37].

2.1.5. Physical Security

Physical security is often overlooked in IoT systems but is equally important. Physical access to devices should be restricted through secure enclosures, locks, and surveillance systems. Additionally, authentication mechanisms like biometrics or smart cards can be implemented to ensure that only authorized personnel can interact with IoT devices physically. Regular audits and physical access point monitoring can help identify and mitigate potential security breaches [38–41].

2.1.6. Lifecycle Management

The entire lifecycle of an IoT device, from development through deployment to disposal, must be carefully addressed to ensure comprehensive security. Secure device provisioning ensures that devices are securely initialized, configured, and deployed. Monitoring mechanisms should be in place to detect anomalies, unauthorized access attempts, or abnormal behavior of IoT devices. At the end of their lifecycle, devices should be decommissioned securely, ensuring that sensitive data are erased and the device cannot be repurposed maliciously.

The fundamentals of IoT security encompass a range of measures to protect against threats and vulnerabilities. Authentication and authorization, secure communication, device and firmware security, data protection and privacy, physical security, and lifecycle management are essential components. By integrating these fundamental principles into the design, development, and deployment of IoT systems, a secure and trustworthy environment can be established, enabling the full potential of the Internet of Things while minimizing risks to individuals, organizations, and critical infrastructure.

2.2. Distributed Reinforcement Learning

Distributed Reinforcement Learning refers to applying reinforcement learning algorithms in a distributed or decentralized setting. In traditional reinforcement learning, an agent interacts with an environment, receiving feedback through rewards or penalties. It uses this feedback to learn a policy that maximizes its cumulative reward over time. The agent's goal is to find an optimal policy that allows it to make the best decisions in the given environment [42].

In a Distributed Reinforcement Learning setup, multiple agents concurrently collaborate and learn from their interactions with the environment. Each agent operates in its environment and learns its policy independently. These agents communicate and share information to enhance their learning process and collectively improve their performance [43]. There are several advantages to using Distributed Reinforcement Learning, summarized as follows.

- *Speed and scalability:* Distributed Reinforcement Learning allows multiple agents to interact with their respective environments in parallel, which can significantly speed up the learning process, especially for complex tasks and large-scale environments.
- *Exploration efficiency:* When agents share knowledge, they can collectively explore the state-action space more effectively, leading to better learning outcomes and more comprehensive policy discovery.
- *Robustness:* By having multiple agents learning in parallel, the system becomes more resilient to failures or noise in individual agents' experiences.
- *Resource utilization:* Distributed Reinforcement Learning can better use computational resources by distributing the learning process across multiple machines or processors.

However, Distributed Reinforcement Learning also introduces new challenges, such as communication overhead, coordination between agents, and handling nonstationarity in the learning environment. Researchers and practitioners often use various techniques and architectures to implement distributed reinforcement learning systems, such as asynchronous methods like Asynchronous Advantage Actor-Critic (A3C) and distributed Deep Deterministic Policy Gradient (DDPG). These approaches help enable efficient collaboration and knowledge sharing among agents to improve performance and learning speed.

3. Distributed Reinforcement Learning for IoT Security

Traditional IoT security solutions struggle to keep pace with IoT systems' dynamic nature, exposing vulnerabilities to ever-evolving threats [16]. There is a growing interest in exploring the potential of Distributed Reinforcement Learning techniques to address these challenges and enhance IoT security.

This section explores the applications of Distributed Reinforcement Learning in IoT security and its potential to transform the protection mechanisms for networked IoT devices. This section examines the fundamental principles underlying IoT security, the unique challenges it presents, and the limitations of traditional security approaches.

3.1. IoT Security in Distributed Networks

The challenges posed by IoT devices' communication protocols and operating systems are significant. The increased attack surface resulting from this diversity may allow enemies to exploit potential vulnerabilities. Implementing centralized security measures in IoT networks is challenging because these networks are decentralized, spanning various contexts and locations.

Figure 1 illustrates the framework for Distributed Reinforcement Learning-based security enforcement in a distributed IoT ecosystem. This figure provides a visual overview of how Distributed Reinforcement Learning is applied to enhance security measures across IoT devices. The framework emphasizes the role of multiple agents learning and adapting to potential threats in real time, facilitating a dynamic and responsive security posture. It showcases how Distributed Reinforcement Learning can manage an IoT ecosystem's complex and varied security requirements, ensuring robust protection against evolving cyber threats.

The security measures designed for traditional IT networks are often inadequate in effectively addressing the distinct security challenges of implementing IoT in distributed networks. The dynamic nature of IoT ecosystems challenges the effectiveness of static security solutions that rely on predetermined rules and configurations. These solutions are not easily adaptable to changing circumstances. It is crucial to prioritize implementing security systems that can continuously adapt and safeguard against emerging threats. This is especially important due to the frequent entry and exit of devices within the network.

Recently, there has been a significant increase in the use of Deep Reinforcement Learning approaches to tackle the security challenges in the Internet of Things. Distributed Reinforcement Learning, a specialized branch of machine learning, enables acquiring knowledge and making informed decisions within complex and constantly changing environments. Utilizing Distributed Reinforcement Learning methodologies can allow the development of adaptive and intelligent security systems to respond to evolving threats in real-time effectively.

3.2. IoT Security in Heterogeneous Networks

Heterogeneous networks, which consist of diverse devices, protocols, and communication technologies, pose unique challenges for IoT security. This section will delve into the technical aspects of IoT security in heterogeneous networks, exploring potential vulnerabilities and advanced security measures to safeguard against threats.

3.2.1. Challenges of IoT Security in Heterogeneous Networks

IoT security in heterogeneous networks is a critical concern due to the diverse nature of devices, protocols, and communication technologies. Heterogeneous networks are characterized by the coexistence of various communication technologies, such as Wi-Fi, Bluetooth, Zigbee, cellular networks, and more, which are often used to connect IoT devices.

Securing IoT devices and data in such environments is challenging because different devices may have varying computational power, memory, and security features. Key security challenges for IoT security in heterogeneous networks include the following:

- *Diverse devices and protocols:* Heterogeneous networks comprise a wide range of IoT devices, each with different hardware capabilities and software architectures. These devices may run on various protocols such as Wi-Fi, Bluetooth, Zigbee, LoRaWAN, and cellular networks, challenging enforcing a unified security framework. Security mechanisms and encryption standards should be adaptable to accom-

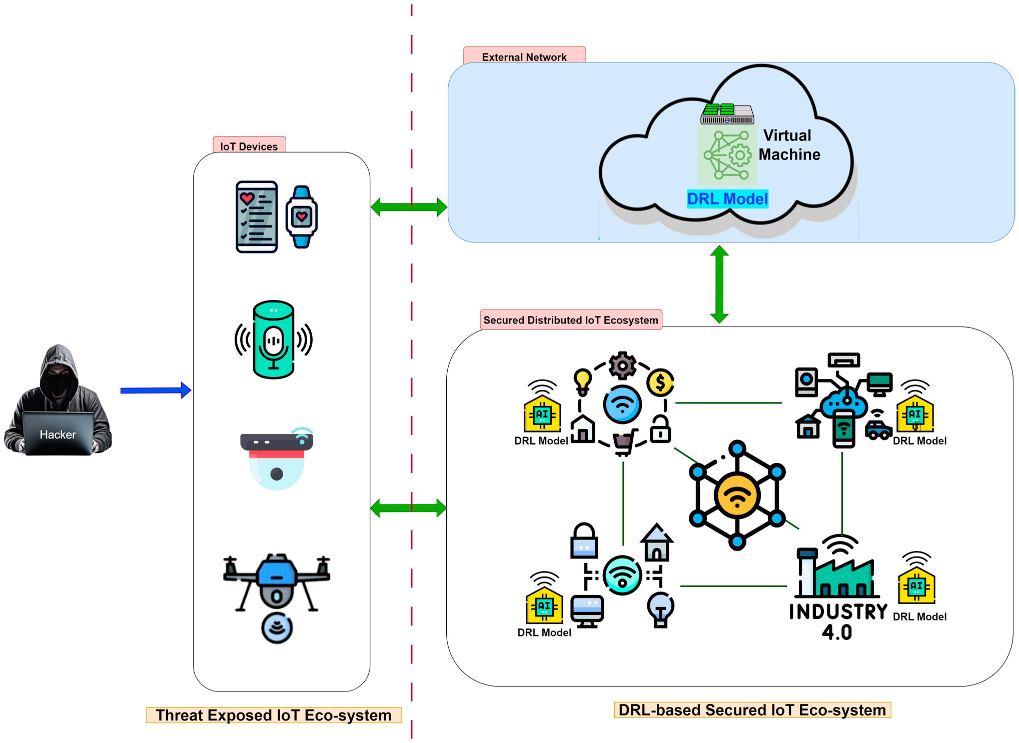

Figure 1: Distributed Reinforcement Learning-based security enforcement in a distributed IoT ecosystem.

modate these devices' varying capabilities and constraints.

- *Interoperability issues:* IoT devices in heterogeneous networks often come from different manufacturers and vendors, leading to interoperability challenges. Integrating devices with varying implementations of security can create weak points in the network. Ensuring seamless communication and robust security requires standardized protocols and strong authentication methods.
- *Scalability:* As the number of IoT devices in a network grows, the complexity of security management increases exponentially. Heterogeneous networks, by their nature, are likely to contain many interconnected devices, elevating the potential attack surface. Traditional security approaches may not scale effectively, necessitating the adoption of distributed security mechanisms that can handle large-scale deployments.
- *Resource constraints:* Many IoT devices in heterogeneous networks are resource constrained in terms of processing power, memory, and energy. Implementing complex security measures on such devices can be challenging and may affect their primary functionalities. Developing lightweight security protocols and efficient encryption algorithms is essential to balance security and resource consumption.

Table 2 compares various Distributed Reinforcement Learning models and their applications in different IoT scenarios. The table includes critical metrics such as performance, security issues addressed, optimization strategies employed, and the specific IoT applications targeted by each Distributed Reinforcement Learning model. It highlights the effectiveness of Distributed Reinforcement Learning approaches in tackling challenges like energy consumption, data tampering, unauthorized access, and intrusion detection across diverse IoT environments. Including metrics like energy efficiency, accuracy in anomaly detection, and security robustness offers a comprehensive overview of the strengths and potential areas for improvement in Distributed Reinforcement Learning-based security systems for IoT.

3.2.2. Security Threats in Heterogeneous IoT Networks

- *Device spoofing and identity misrepresentation:* Due to the diverse nature of IoT devices, attackers can exploit the lack of proper authentication and authorization mechanisms to impersonate devices or manipulate their identities. This can lead to unauthorized access, data breaches, and unauthorized control over critical systems.
- *Man-in-the-Middle (MITM) attacks:* Heterogeneous networks can be susceptible to MITM attacks, where

an attacker intercepts and alters the communication between IoT devices. Weak encryption and unauthenticated communication channels can facilitate MITM attacks, leading to data manipulation, injection of malicious payloads, and unauthorized access.

- *DoS and Distributed Denial-of-Service (DDoS) attacks:* IoT devices with limited resources can be exploited to launch DoS and DDoS attacks. Attackers can overwhelm devices with massive traffic, rendering them non-functional and disrupting critical services.
- *Firmware and Software Vulnerabilities:* IoT devices often run on firmware or software that may contain security vulnerabilities. Attackers can exploit these weaknesses to gain unauthorized access, compromise the integrity of devices, or launch attacks on other parts of the network.

3.2.3. Advanced Security Measures

- *Secure boot and device attestation:* Implementing secure boot mechanisms ensure that only authorized and unmodified firmware is executed on IoT devices. Device attestation enables the verification of device identity, ensuring that only legitimate devices are allowed to join the network.
- *Mutual authentication:* Enforcing mutual authentication between devices and gateways ensures that both parties can verify each other's identities before establishing communication. This prevents device spoofing and unauthorized access.
- *End-to-end (E2E) encryption:* Utilizing strong end-to-end encryption protocols, such as TLS (Transport Layer Security), safeguards data transmitted between IoT devices and cloud servers. This prevents eavesdropping and data tampering during transit.
- *Intrusion detection and anomaly detection:* Implementing intrusion detection systems (IDS) and anomaly detection mechanisms helps identify suspicious activities and potential security breaches. Machine learning algorithms can detect abnormal behavior patterns within the network.
- *Over-the-Air (OTA) updates:* OTA updates allow IoT devices to receive security patches and software updates remotely, ensuring that devices stay protected against newly discovered vulnerabilities.

3.3. A Typical Reinforcement Learning Solution

To address the significant security challenges posed by heterogeneous and distributed IoT networks, this paper introduces an efficient Distributed Reinforcement Learning framework. This framework uses decentralized decision-making and adaptive learning techniques to secure IoT environments dynamically. The proposed solution effectively mitigates threats, while maintaining scalability and optimizing resource usage by employing multiple Distributed Reinforcement Learning agents that operate collaboratively.

3.3.1. Key Framework Components

The key components of the proposed framework are described below:

- **State space:** This component identifies critical features of IoT network traffic that the agents use to make decisions. These features include:
  - Packet size, which indicates the amount of data being transmitted.
  - Protocol type helps understand the type of communication taking place.
  - Source and destination IP addresses, which reveal the origins and targets of the network traffic.
  - Time-based traffic patterns, which provide insights into the behavior of network usage over time.
- **Action space:** This defines the set of possible actions that agents can take to mitigate security threats, such as:
  - Classifying network traffic as either normal or abnormal.
  - Initiating containment measures to isolate potentially compromised devices.
  - Alerting network administrators for further investigation and action.
- **Reward function:** The framework employs a reward system to guide learning by encouraging desirable and discouraging undesirable actions. For instance:
  - Correctly identifying a security threat earns a positive reward.
  - Misclassifying traffic as false positives or negatives results in a penalty.
  - Delays responding to a threat are also penalized to ensure timely action.
- **Learning algorithm:** The framework uses an adaptive learning method to help each agent improve its decision-making over time. By analyzing past actions and outcomes, agents refine their strategies to better detect and mitigate security threats in real-world IoT scenarios.

Table 2: Overview of Distributed Reinforcement Learning models and their applications in enhancing IoT security. **Note:** DRL stands for Distributed Reinforcement Learning in this table.

IoT Application	DRL Model	Performance Metrics	Security Issues	Optimization Strategy	Security Solutions
Smart home monitoring	A3C	Energy consumption, Privacy leakage & Decentralized training	Homomorphic encryption, Intrusion detection accuracy & Data tampering	FL, Secure FL & Communication overhead	Unauthorized Access, Model compression & Access control
Industrial control systems	DDPG	Control latency, DoS attacks & Parameter server architecture	Anomaly detection algorithms, System stability & Manipulation of control signals	Network partitioning, Control signal encryption & Resource utilization	Data interception, Prioritized experience replay & Network segmentation
Autonomous vehicles	DQN	Decision-making speed, Sensor data spoofing & Experience replay	Sensor data authentication, Collision avoidance accuracy & Communication hijacking	Prioritized exploration, V2Xsecurity & Navigation efficiency	Malicious obstacle insertion, Ensemble Learning & Multi-Sensor Fusion
Healthcare wearables	PG	Vital sign monitoring accuracy, Health data interception & Actor-critic architecture	End-to-end encryption, Real-time monitoring & Privacy violation	Parameter sharing, Differential privacy techniques & Energy efficiency	Device malfunction, On-device learning & Device authentication
Smart grid management	PPO	Energy distribution efficiency, False data injection & multi-agent learning	Secure state estimation, Load balancing & Unauthorized access	Reinforcement Learning with Expert Knowledge, Anomaly detection in energy flow & Grid stability	Demand manipulation, Dynamic rewards & Authentication protocols
Smart agriculture	MADDPG	Crop yield, Sensor data tampering & Decentralized multi-agent learning	Data integrity verification, Resource utilization & Unauthorized access	Resource allocation policies, Secure multi-agent communication & Pest detection accuracy	Data privacy, Cross-agent knowledge sharing & Encrypted sensor data
Environmental monitoring	TRPO	Data accuracy, sensor spoofing & Sensor fusion techniques	Data authentication, Real-time monitoring & Data manipulation	Online learning, Tamper-proof sensor design & Energy efficiency	Communication interception, Energy-efficient Communication protocols & IDS
Energy-efficient buildings	A2C	Energy consumption, Unauthorized control & Hierarchical multi-Level learning	Device authentication, temperature regulation & Energy theft	Energy-efficient RL, Anomaly detection in energy Usage & Occupancy prediction	Privacy breach, Stochastic policies & Role-based access control

3.3.2. Typical Implementation and Workflow

1. **Data collection and preprocessing:** IoT network traffic data (e.g., from CICIDS2017 [44] and IoT-23 [45] datasets) is preprocessed to extract relevant features for state representation.
2. **Agent deployment:** Each Distributed Reinforcement Learning agent monitors a subset of the network, operating autonomously to identify and respond to anomalies in real-time.
3. **Collaborative learning:** Agents periodically share learning outcomes to enhance threat detection and minimize redundant actions.
4. **Action execution:** Upon detecting an anomaly, an agent executes the optimal action determined by its policy (e.g., isolate a device, notify administrators).

3.3.3. Advantages of the Distributed Reinforcement Learning framework

- **Scalability:** The decentralized nature of Distributed Reinforcement Learning allows it to scale seamlessly by adding new devices or network segments.
- **Adaptability:** Agents dynamically adjust to evolving threat patterns, making the system robust against zero-day attacks.
- **Efficiency:** Lightweight agents are tailored for resource-constrained IoT devices, ensuring minimal impact on performance.

4. Datasets and Use Cases for Distributed Reinforcement Learning for IoT Security

This section discusses the most common datasets used in the field of Distributed Reinforcement Learning-based intrusion detection in IoT networks and different use case applications of Distributed Reinforcement Learning in IoT security.

4.1. Commonly Used Datasets

The availability of high-quality datasets is crucial for training and evaluating models [46]. Here are some of the most commonly used datasets:

- **NSL-KDD [47]**: An improved version of the KDD'99 dataset, NSL-KDD addresses some of the inherent issues of its predecessor by removing redundant records and ensuring a balanced class distribution. This dataset is widely used for intrusion detection system (IDS) evaluations.
- **UNSW-NB15 [48]**: This dataset contains nine types of attacks along with normal network activities. It is generated using the IXIA PerfectStorm tool to create real modern normal activities and synthetic contemporary attack behaviors.
- **CICIDS2017 [44]**: This dataset was created by the Canadian Institute for Cybersecurity and includes a comprehensive set of real-world traffic scenarios to generate a dataset for IDS evaluations.
- **IoT-23 [49]**: A dataset of network traffic from IoT devices containing labeled traffic traces from real IoT devices operating in various scenarios, including benign and malicious behavior.

4.2. Use Case: Application of Distributed Reinforcement Learning in IoT Security

4.2.1. Scenario Description

Consider a smart home environment with multiple IoT devices, such as smart thermostats, security cameras, and lighting systems. These devices communicate over a heterogeneous network and are susceptible to various security threats, including unauthorized access, data breaches, and DoS attacks.

4.2.2. Implementation

To strengthen the security of the IoT ecosystem, a Distributed Reinforcement Learning-based intrusion detection system (IDS) that continuously monitors network traffic and detects abnormal behavior.

1. **Environment setup**: The network traffic data, including both normal and malicious activities, is collected from the IoT devices. This data is preprocessed and fed into the Distributed Reinforcement Learning model.
2. **Distributed Reinforcement Learning model**: A Deep Q-Network (DQN) is utilized for the Distributed Reinforcement Learning model. The state space includes features extracted from the network traffic, such as packet size, source and destination IP addresses, and protocol types. The action space consists of possible decisions, such as flagging the traffic as normal or abnormal and triggering appropriate security measures.
3. **Training phase**: The DQN is trained using the preprocessed dataset. The reward function is designed to penalize false positives and false negatives while rewarding correct classifications. This encourages the model to improve its detection accuracy over time.
4. **Deployment**: The trained DQN model is deployed in the smart home network. It monitors real-time traffic and makes decisions based on the learned policies. When an anomaly is detected, it triggers predefined security measures, such as isolating the affected device or alerting the user.

The method of implementing an IDS for improving security in an IoT ecosystem using Distributed Reinforcement Learning is highlighted in Algorithm 1. The first step of the procedure is gathering and preprocessing network traffic data, which covers both benign and malevolent/malicious activity (Environment Setup). After that, the Distributed Reinforcement Learning model, precisely a DQN, is defined using this data. In this model, the action space consists of choices like classifying traffic as normal or abnormal, and the state space comprises attributes taken from the network traffic. The preprocessed dataset is used to train the DQN during the training phase. A reward function that penalizes false positives and false negatives incentivizes accurate classifications. The DQN model is deployed in the IoT network after it has been trained.

4.3. Mathematical Modeling of Distributed Reinforcement Learning-Based IDS

This subsection introduces a straightforward mathematical model of Distributed Reinforcement Learning in the context of IDS.

Algorithm 1 Implementation of Distributed Reinforcement Learning-based IDS in IoT security.

```
Input: Preprocessed IoT network traffic data $D = \{d_1, d_2, ..., d_N\}$
Output: Trained Deep Q-Network (DQN) model for anomaly detection
procedure TRAINING PHASE
    Initialize IoT environment with state space S, action space A, and reward function R.
    Configure DQN model with state representation $s \in$ S and possible actions $a \in$ A.
    for each training episode $e \in E$ do
        Observe current state $s_t$ from environment
        Select action $a_t$ using an $\epsilon$-greedy policy.
        Execute action $a_t$ and observe reward $r_t$ and next state $s_{t+1}$.
        Update Q-value estimate based on observed reward and next state.
    end for
    Save the trained DQN model.
end procedure
procedure DEPLOYMENT PHASE
    Deploy trained DQN model to monitor real-time IoT network traffic.
    while true do
      Observe real-time state $s_t$ from incoming traffic.
Predict optimal action $a_t$ using the DQN model.
      if $a_t$ indicates an anomaly then
Trigger security measures:
•     Isolate the compromised IoT device.
•     Alert the system administrator.
      else
          Continue monitoring traffic.
      end if
    end while
end procedure
return Deployed anomaly detection model
```

1. **Environment setup**: The state space S represents features extracted from the network traffic:

$$S = \{\text{packet size, source IP, destination IP, protocol type, } \ldots\}$$

The action space A includes possible decisions:

$$A = \{\text{normal, anomalous}\}$$

2. **Distributed Reinforcement Learning model**: Using Q-learning, the Q-value function $Q(s, a)$ represents the expected cumulative reward of taking action a in state s:

$$Q(s,a) = E[\textstyle\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \mid s_0 = s, a_0 = a] \tag{1}$$

where γ is the discount factor.

3. **Training phase**: The reward function R is defined as:

$$R(s, a) = \begin{array}{l} +1 \text{ if correctly classified} \\ \\ -1 \text{ if incorrectly classified} \end{array} \tag{2}$$

The Q-learning update rule is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[R(s,a) + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$$

where α is the learning rate.

4. **Deployment**: The trained DQN model is deployed in the IoT network to monitor real-time traffic. The policy π is defined as:

$$\pi(s) = \operatorname*{argmax}_a Q(s,a) \tag{3}$$

4.4. Complexity Analysis of the Proposed Distributed Reinforcement Learning Scheme

The proposed Distributed Reinforcement Learning framework introduces computational complexity at various stages of its operation. This subsection analyzes the complexity of communication, computation, and resource utilization to comprehensively understand its feasibility for IoT networks.

4.4.1. Computational Complexity

The Distributed Reinforcement Learning agents utilize a DQN to approximate the Q-values for decision-making. The computational complexity per update of the DQN is primarily determined by:

- The size of the input state space (S) depends on the number of features extracted from the IoT traffic data.

- The depth and number of neurons in the neural network, denoted as $O(d \cdot n^2)$, where d is the depth of the network and n is the number of neurons per layer.
- The number of training iterations required for convergence depends on the reward structure and the dynamic nature of the environment.

For typical IoT networks, the framework's computational requirements can be scaled down by optimizing the neural network architecture to match the resource constraints of the devices.

4.4.2. Communication Overhead

Agents periodically share their learning outcomes in the proposed Distributed Reinforcement Learning framework to enhance collaborative threat detection. The communication overhead is determined by:

- The frequency of information exchange among agents.
- The size of the shared information, which includes Q-value updates and aggregated state-action pairs.
- The network topology, where more dispersed networks may experience higher latency and energy consumption during communication.

The communication overhead can be reduced by employing sparse communication schedules and compressing the transmitted data.

4.4.3. Resource Utilization

IoT devices are often constrained by limited computational power, memory, and energy. The proposed Distributed Reinforcement Learning framework addresses these constraints by:

- Deploying lightweight agents that offload computationally intensive tasks to edge or cloud resources when necessary.
- Optimizing the reward function to minimize unnecessary computations, thus conserving energy.
- Implementing decentralized learning to distribute the workload across multiple agents, avoiding bottlenecks.

4.4.4. Scalability and Adaptability

The Distributed Reinforcement Learning framework is inherently scalable due to its decentralized nature. The complexity scales linearly with the number of devices ($O(n)$), as each agent operates independently while collaborating only periodically. This characteristic ensures that the framework remains adaptable to dynamic IoT environments with varying numbers of devices.

4.4.5. Summary

While the proposed Distributed Reinforcement Learning framework introduces certain computational and communication overheads, its design incorporates optimizations to align with the resource constraints of IoT devices. By balancing the trade-offs between complexity and performance, the framework ensures efficient and effective security in heterogeneous IoT networks.

5. Evaluation Metrics and Performance Analysis

Distributed Reinforcement Learning in IoT security requires robust evaluation to ensure effectiveness across diverse and dynamic environments. Comprehensive performance analysis involves multiple dimensions, including security, efficiency, adaptability, and resource utilization. This section expands on these metrics and integrates insights from recent literature to provide a broader perspective.

5.1. IoT Security Metrics

- Detection accuracy: Accurate identification of security threats is a fundamental metric. Studies, such as Wang et al. [50] and Li et al. [51], emphasize optimizing threat detection accuracy to mitigate risks like data breaches and unauthorized access.
- False positive and negative rates: Minimizing both false alarms and missed threats is critical [52]. Kumar and Singh [53] discuss reinforcement learning models tailored to reduce these rates in DDoS attack prevention.
- Response time: The ability to detect and respond to threats in real-time is vital for IoT applications [54]. Mishra et al. [55] highlight the significance of response time in security-critical systems like medical IoT.
- Resilience against adversarial attacks: Robustness to adversarial inputs ensures consistent performance under malicious conditions [56]. Benaddi et al. [57] explore methods integrating Generative Adversarial Networks (GANs) to enhance resilience.

5.2. Efficiency and Resource Utilization Metrics

- Energy consumption: IoT devices often operate under resource constraints, requiring energy-efficient DRL algorithms [58]. For instance, Louati et al. [59] discuss strategies to optimize energy utilization in distributed intrusion detection systems.

- Scalability: The ability to scale across heterogeneous networks is essential for DRL frameworks. Studies such as Abou El Houda et al. [60] demonstrate scalability in jamming attack mitigation using federated DRL models.

5.3. General Performance Metrics

- Learning efficiency: Efficient training processes are crucial for real-time adaptability. For example, Diro et al. [61] analyze Distributed Reinforcement Learning techniques that prioritize faster convergence in industrial IoT scenarios.
- Interoperability: Compatibility with diverse IoT devices and protocols enhances the applicability of DRL solutions. Studies like Bikos et al. [62] provide insights into interoperability challenges and solutions.
- Long-term adaptability: The ability to adapt to evolving IoT environments ensures sustained performance [63]. Feng et al. [64] discuss DRL strategies designed for dynamic threat landscapes.

In summary, the evaluation of Distributed Reinforcement Learning in IoT encompasses a broad spectrum of metrics, reflecting the multifaceted challenges and requirements of IoT systems. From security to operational efficiency, these metrics provide a comprehensive view of how well Distributed Reinforcement Learning algorithms perform and where improvements may be needed. As IoT continues to grow in scope and complexity, the role of Distributed Reinforcement Learning in managing and securing these networks will likely become even more significant, making these evaluation metrics the more essential.

5.4. Comparisons of Related State-of-the-art Models

Table 3 provides a comprehensive comparison of the proposed Distributed Reinforcement Learning framework with recent and relevant works in the field. These works highlight a variety of methodologies, ranging from deep learning and reinforcement learning to hybrid approaches such as federated learning and GAN-integrated models. This section discusses the strengths and limitations of these methodologies and positions the proposed DRL framework as a robust solution for IoT security challenges.

Table 3: Comprehensive comparison of state-of-the-art DRL frameworks. **Note:** DRL, RL, DL, and FDR stand for Distributed Reinforcement Learning, Reinforcement Learning, Deep Learning, and Federated Deep Reinforcement, respectively, in this table.

Ref.	Methodology	Strengths	Limitations
[50]	DRL for Age of Information Minimization	Optimizes real-time communication; minimizes age of information	Focuses on communication latency; less emphasis on threat detection
[53]	RL for DDoS Detection and Prevention in Edge IoT	Effective against DDoS attacks; reduces network traffic overhead	Limited to DDoS scenarios; does not address other IoT threats
[65]	DL for IoT Attack Mitigation	High accuracy in attack detection; optimizes learning parameters	Computationally intensive; lacks focus on resource-constrained IoT devices
[60]	FDR Learning for Jamming Attack Mitigation	Efficient in distributed environments; preserves privacy	Applicable primarily to jamming attacks; requires high computational resources
[55]	Cognitive DRL for Medical Cyber-Physical Systems	Adaptable to dynamic environments; improves medical system security	Limited scalability; specific to medical IoT applications
[66]	DRL for Industrial IoT Security	Proactive threat mitigation; robust to evolving attacks	High computational overhead; may not scale well in large networks
[62]	RL for Anomaly Detection in IoT with DLT	Integrates DLT for secure ledgering; effective for anomaly detection	High resource consumption due to DLT integration
[67]	Distributed DL for IoT Attack Detection	Distributed architecture; high detection accuracy	Communication overhead in distributed systems
[59]	Multi-Agent RL for Big Data IoT Intrusion Detection	Decentralized intrusion detection; scales with big data networks	High training complexity; requires substantial coordination
[57]	DRL with GAN for Anomaly Detection	Combines GAN and DRL for robust anomaly detection	Complexity due to GAN integration; high computational cost

Table 3: *Cont.*

Ref.	Methodology	Strengths	Limitations
[61]	Distributed DL for IoT Attack Detection	High detection accuracy; effective in large-scale IoT networks	Limited adaptability to emerging threats
[68]	Trust-Driven Reinforcement Selection Strategy in FL	Enhances trust in federated learning models for IoT devices	Limited scalability in highly dynamic IoT environments
[69]	Federated DRL for Secure IoT Data Sharing	Preserves data privacy; enhances collaboration in distributed IoT	High computational complexity; requires extensive coordination
[64]	DRL for Security Defense in IoT	Proactive defense strategy; reduces response time to attacks	Computational overhead; requires fine-tuning for diverse IoT applications

Several works, such as [50,69], leverage distributed reinforcement learning and federated learning, respectively, to address the dynamic and heterogeneous nature of IoT systems. While these approaches are well-suited for decentralized environments, their primary focus is either on optimizing communication efficiency or ensuring data privacy, with limited emphasis on comprehensive threat detection across diverse IoT use cases. For example, ref. [50] minimizes the age of information in real-time systems but does not address security-specific challenges. Similarly, ref. [69] enhances secure data sharing but faces scalability issues due to the high computational complexity of federated learning.

Other works, such as [53,59], have applied reinforcement learning and multiagent reinforcement learning to detect and prevent specific threats like DDoS attacks and intrusions. These frameworks demonstrate the efficacy of reinforcement learning in achieving adaptive security. However, ref. [53] is limited in scope to DDoS scenarios, and [59] encounters high training complexity when scaling to big data networks.

Hybrid approaches, such as the combination of GANs and DRL in [57], provide robust anomaly detection mechanisms but are hindered by their computational cost and implementation complexity. Furthermore, works like [68] focus on trust-driven strategies in federated learning, which enhances model reliability but may not scale effectively in dynamic IoT environments.

The proposed DRL framework addresses several of these limitations by combining the adaptability of reinforcement learning with decentralized decision-making tailored for IoT systems. Unlike centralized or computationally intensive approaches, the proposed framework is designed to operate efficiently in resource-constrained environments while maintaining scalability and robustness to evolving threats. By focusing on real-time threat detection, dynamic reward optimization, and collaboration among agents, the proposed framework offers a holistic solution for securing heterogeneous IoT networks.

6. Future Directions and Open Challenges

The future of Distributed Reinforcement Learning for IoT security in distributed and heterogeneous networks holds significant potential for transformative research. By addressing the following research directions, researchers can pave the way for more robust, adaptive, and secure IoT systems capable of defending against emerging security threats in a dynamic and interconnected world.

6.1. Hybrid Reinforcement Learning Models

A promising approach to enhancing IoT security in distributed and heterogeneous networks is the development of hybrid reinforcement learning models. These models seek to strike a balance between centralized and decentralized approaches, capitalizing on the strengths of each. Centralized models can leverage global information, enabling more informed and optimized decision-making. On the other hand, decentralized models offer improved scalability and reduced communication overhead by allowing individual IoT devices to make independent decisions based on their local observations. By combining the benefits of both approaches, researchers can design reinforcement learning algorithms that are more efficient and more effective in addressing the unique security challenges of IoT networks. These hybrid models could lead to adaptive and intelligent security measures that respond dynamically to evolving threats while ensuring seamless coordination and cooperation across the distributed IoT ecosystem.

Moreover, they have the potential to optimize resource allocation and energy consumption, making them well-suited for resource-constrained IoT devices. However, achieving the right balance and fine-tuning the parameters of such hybrid models pose considerable research challenges, necessitating innovative techniques and algorithm design to ensure their successful deployment in real-world IoT security scenarios. As the IoT landscape expands and diversifies, exploring hybrid reinforcement learning models holds great promise for revolutionizing IoT security and fortifying interconnected devices against emerging threats.

6.2. Federated Learning for Privacy Preservation

In IoT security, a compelling area of exploration lies in applying federated learning techniques to distributed reinforcement learning. Federated learning presents a robust approach where IoT devices can collaboratively train a shared model while keeping their data localized, ensuring privacy preservation. This is especially relevant in IoT scenarios where data privacy is paramount, as it mitigates the risks of sensitive information exposure during the learning process.

By enabling IoT devices to learn from collective experiences without sharing raw data, federated learning empowers the creation of robust and accurate models while respecting user privacy and adhering to data protection regulations. In this context, federated learning can be harnessed to develop adaptive and localized security models for individual IoT devices, capturing device-specific nuances and addressing their unique security requirements.

Furthermore, federated learning fosters efficient model updates across diverse IoT devices, promoting scalability and reducing the communication burden in large-scale distributed networks. However, implementing federated learning in the context of distributed reinforcement learning for IoT security presents technical challenges, including communication optimization, model aggregation, and ensuring convergence with diverse data distributions and device capabilities. As research progresses, exploring the potential of federated learning in IoT security promises to unlock privacy-aware and decentralized security solutions that safeguards the IoT ecosystem while respecting the privacy rights of users.

6.3. Adversarial Robustness

In bolstering IoT security in distributed and heterogeneous networks, focusing on enhancing the robustness of distributed reinforcement learning models against adversarial attacks is essential. Adversarial attacks on IoT devices pose significant threats, potentially leading to severe consequences, making the resilience of reinforcement learning models crucial. To address this, researchers can investigate various methods to fortify the models against such attacks.

Adversarial training is a prominent technique that involves exposing the reinforcement learning model to carefully crafted adversarial examples during training. By incorporating these adversarial examples, the model learns to recognize and defend against potential attack scenarios, making it more robust when faced with real-world adversarial attempts.

Secure aggregation is another approach to safeguard the reinforcement learning process in distributed IoT networks. The model updates from different devices are aggregated in secure aggregation while maintaining data privacy and confidentiality. This ensures that malicious devices cannot manipulate the aggregation process to introduce adversarial perturbations, thereby protecting the integrity of the collective model.

Furthermore, researchers can explore ensemble techniques, combining multiple reinforcement learning models to increase robustness. Using an ensemble of models, the system can collectively make more reliable decisions, minimizing the impact of adversarial attacks.

Moreover, investigating transfer learning methods is valuable. Transfer learning allows models to leverage knowledge gained from previous experiences in similar domains or tasks. By pre-training models on a large dataset from a different but related domain and then fine-tuning them on the target IoT security task, the models can inherit some robustness from the initial training, providing a head start in defending against adversarial attacks.

To evaluate the effectiveness of these methods, researchers can conduct extensive testing in realistic IoT environments with diverse device configurations, network topologies, and attack scenarios. Benchmarking the performance against various adversarial attacks will be essential to assess the models' resilience and identify potential vulnerabilities needing further reinforcement.

In addressing the challenge of enhancing the robustness of distributed reinforcement learning models against adversarial attacks in IoT security, researchers must also take into account the computational overhead introduced by these defensive measures. Striking a balance between robustness and efficiency is critical, as resource-constrained IoT devices may face limitations in processing power and energy consumption.

Investigating methods to fortify Distributed Reinforcement Learning models against adversarial attacks is vital for ensuring the security and resilience of IoT devices in distributed and heterogeneous networks. Adversarial training, secure aggregation, ensembling, and transfer learning are among the potential techniques that researchers can explore. Rigorous evaluation and optimization of these approaches are essential to develop effective and efficient defenses against adversarial threats in the evolving landscape of IoT security.

6.4. Resource-Efficient Distributed Reinforcement Learning

It is imperative to develop resource-efficient Distributed Reinforcement Learning algorithms that can operate effectively on IoT devices with limited computational capabilities and memory to enhance IoT security in resource-constrained environments. To achieve this, researchers must address the challenge of reducing Distributed Reinforcement Learning models' computational and memory requirements while ensuring their security effectiveness.

One effective method for achieving resource efficiency is through model compression. This technique involves reducing the size and complexity of Distributed Reinforcement Learning models while maintaining their high performance. Techniques such as quantization, pruning, and knowledge distillation can be investigated to develop compact yet precise models ideal for deployment on IoT devices with limited resources.

Furthermore, designing specialized architectures tailored to the constraints of IoT devices can significantly enhance resource efficiency. Customizing neural network architectures and algorithms can help reduce Distributed Reinforcement Learning models' computational burden, memory footprint, and energy consumption, making them more amenable to resource-constrained IoT environments.

On-device federated learning is another promising avenue to improve resource efficiency in distributed reinforcement learning. By enabling IoT devices to participate in collaborative learning without sharing raw data, on-device federated learning minimizes communication overhead. It reduces the need for central processing, making it suitable for resource-constrained settings.

Researchers can explore techniques like transfer learning and incremental learning to ensure security effectiveness while optimizing resource usage. By leveraging pre-trained models or updating the model incrementally based on new data, the Distributed Reinforcement Learning algorithms can adapt and improve over time without requiring large-scale retraining, thereby conserving computational resources.

Moreover, exploiting domain knowledge and contextual information can lead to more efficient Distributed Reinforcement Learning algorithms. Incorporating prior knowledge about the task or environment can guide the learning process, which reduces the need for extensive exploration and accelerating convergence, particularly crucial for resource-constrained IoT devices.

To evaluate the performance of resource-efficient Distributed Reinforcement Learning algorithms, researchers can conduct thorough testing on a range of IoT devices with varying resource capacities and network conditions. Benchmarking against traditional Distributed Reinforcement Learning models and assessing security effectiveness will provide insights into the trade-offs between resource efficiency and security.

Collaboration with hardware and system-level experts is essential to optimize the implementation of resource-efficient Distributed Reinforcement Learning algorithms on IoT devices. By leveraging hardware accelerators, hardware-aware optimization, and platform-specific optimizations, the Distributed Reinforcement Learning models can make the most of the available resources while delivering robust security outcomes.

Ultimately, the development of resource-efficient Distributed Reinforcement Learning algorithms for IoT security is a multi-faceted endeavor that demands a deep understanding of both reinforcement learning techniques and the resource constraints of IoT devices. By addressing these challenges, researchers can empower a broader range of IoT devices to adopt sophisticated Distributed Reinforcement Learning-based security solutions, fortifying the IoT ecosystem against potential threats while conserving the use of valuable resources.

6.5. Transfer Learning in Heterogeneous Networks

Transfer learning aims to leverage knowledge gained from one IoT device or network segment and apply it to improve the learning process in other segments, leading to faster convergence and better generalization. Individual devices may have varying capabilities, data distributions, and environmental conditions in heterogeneous IoT networks. This diversity poses challenges for traditional Distributed Reinforcement Learning algorithms, as they may struggle to adapt and generalize across different devices effectively. Transfer learning addresses this issue by enabling knowledge transfer between devices, allowing them to learn from experiences on other devices with similar or related tasks.

One approach to transfer learning in Distributed Reinforcement Learning is pretraining a model on a large dataset from a source device or network segment, which possesses abundant data or computational resources. This pre-trained model can be fine-tuned on the target device or segment, leveraging the initial learning to enhance convergence speed and overall performance. This process allows devices with limited data or computational resources to benefit from the knowledge accumulated from other devices, making learning more efficient and effective.

Additionally, transfer learning can facilitate the adaptation of Distributed Reinforcement Learning models to changes in the IoT environment. As IoT networks are dynamic and evolving, the knowledge acquired from past experiences on one device can be transferred to adapt the model to new scenarios, enhancing its generalization capabilities.

However, applying transfer learning in heterogeneous IoT networks requires addressing various challenges. For instance, domain adaptation is crucial to account for variations in data distributions across different devices. Techniques like domain adaptation and domain generalization must be explored to ensure that knowledge transfer is effective across diverse IoT environments.

Furthermore, privacy preservation is a significant concern when transferring knowledge between devices. Privacy-preserving transfer learning methods should be considered to safeguard sensitive information while enabling knowledge exchange.

Researchers can conduct comprehensive experiments across diverse devices and environments to evaluate the potential of transfer learning in distributed reinforcement learning for heterogeneous IoT networks. By evaluating the performance of models with and without transfer learning and comparing them against conventional approaches in Deep Reinforcement Learning, we can gain insights into the advantages and limitations of transfer learning in this specific context.

Overall, the study of transfer learning techniques in heterogeneous IoT networks has the potential to revolutionize Distributed Reinforcement Learning, enabling devices to benefit from shared knowledge and accelerate learning. By addressing challenges related to domain adaptation, privacy preservation, and performance evaluation, researchers can unlock the full potential of transfer learning in empowering IoT devices to collaborate, adapt, and thrive in the diverse and dynamic IoT ecosystem.

6.6. Edge Computing Integration

With its localized processing and decision-making capabilities, edge computing can significantly contribute to reducing latency and improving responsiveness in IoT security applications. Researchers can optimize Distributed Reinforcement Learning performance by investigating how edge nodes can collaborate with central servers while upholding a robust security framework.

In the proposed integration, edge nodes are pivotal in executing certain Distributed Reinforcement Learning computations locally, thereby offloading the central servers and reducing communication latency. This distributed approach empowers edge nodes to make swift and autonomous decisions, enhancing the real-time responsiveness of IoT security systems. Edge computing also addresses concerns related to data privacy and bandwidth consumption. By processing sensitive data locally on edge nodes, there is a reduced need to transmit raw data to central servers, minimizing the risk of data exposure and conserving valuable network resources.

To fully leverage the benefits of edge computing, researchers must investigate efficient algorithms for task allocation and workload distribution between edge nodes and central servers.

Dynamic task allocation mechanisms can adapt to changing workloads and network conditions, ensuring optimal resource utilization and scalability in large-scale IoT deployments. Moreover, the integration of edge computing and Distributed Reinforcement Learning presents opportunities to address security challenges more effectively. For instance, edge nodes can serve as a first line of defense, locally detecting and mitigating common security threats before escalating critical issues to the central servers for further analysis and response.

However, several considerations must be addressed to ensure the effectiveness and security of the integrated approach. Researchers should develop robust authentication and access control mechanisms to prevent unauthorized access to edge nodes and maintain the integrity of the Distributed Reinforcement Learning system.

Furthermore, the potential trade-offs between computation and communication overhead should be carefully evaluated. While edge computing can reduce latency, there might be increased computational burdens on edge nodes, necessitating algorithms, and resource allocation optimization to strike the right balance.

In the investigation, real-world experimentation and validation are crucial to assess the benefits and limitations of the integrated system. Testbeds involving heterogeneous IoT devices, various edge computing configurations, and diverse security scenarios will provide insights into the practical feasibility and performance improvements achieved.

Additionally, researchers must consider the impact of the distributed nature of edge computing on model synchronization and update mechanisms. Techniques like federated learning can be explored to enable seamless collaboration and communication between edge nodes and central servers, ensuring a synchronized and up-to-date Distributed Reinforcement Learning model across the IoT network.

Exploring the integration of edge computing with distributed reinforcement learning holds great promise in advancing IoT security. Leveraging edge nodes for localized processing and decision-making can improve responsiveness and reduce latency. Investigating collaborative approaches between edge nodes and central servers will optimize the performance of Distributed Reinforcement Learning while maintaining a robust security framework. Realizing the full potential of this integrated approach requires addressing technical challenges related to task allocation, workload distribution, security mechanisms, and model synchronization. Researchers can pave the way for efficient, secure, and responsive IoT security systems by conducting comprehensive research and validation.

6.7. Multi-Objective Reinforcement Learning

Multi-objective reinforcement learning presents a nuanced approach to managing the intricate balance of various security needs in distributed IoT networks. This method allows for the simultaneous pursuit of multiple goals, which is essential in the complex landscape of IoT security. For instance, besides minimizing false positives and maximizing detection accuracy, this approach can also focus on reducing energy consumption, a critical factor in IoT devices often constrained by battery life. Additionally, it can be tailored to prioritize swift responses to security threats, ensuring prompt mitigation and minimizing disruptions to network operations.

Moreover, multi-objective reinforcement learning can adapt to different IoT environments' specific requirements and constraints. For example, in a scenario with limited network bandwidth, the model might prioritize reducing the amount of data transmitted while maintaining a high-security vigilance level. The model could focus on rapidly adapting to changes in more dynamic environments, such as new devices joining the network or varying user behavior patterns. By addressing these varied objectives, multi-objective reinforcement learning enables the creation of a more holistic and effective security strategy. This strategy not only addresses the immediate security concerns but also considers the long-term sustainability and efficiency of the IoT network, making it a highly adaptable and comprehensive solution for IoT security challenges.

6.8. Real-World Deployment and Evaluation

The importance of thorough real-world testing and evaluation of Distributed Reinforcement Learning models in the context of IoT security must be emphasized. This practical assessment is crucial to validating the models' efficacy and adaptability. The goal of such deployments should be to rigorously test Distributed Reinforcement Learning models under diverse and often challenging conditions characteristic of IoT environments. This includes evaluating the models' performance in the face of network latency, a common issue in IoT networks, which can significantly impact the speed and reliability of data transmission and, consequently, the responsiveness of the Distributed Reinforcement Learning models. Additionally, the diverse nature of IoT devices, with varying computational capabilities, operating systems, and communication protocols, adds another layer of complexity. It is crucial to ensure that Distributed Reinforcement Learning models can operate efficiently across device heterogeneity, adapting their learning and decision-making processes accordingly. Furthermore, IoT networks often undergo dynamic topology changes, with devices frequently joining and leaving the network. This fluid nature requires Distributed Reinforcement Learning models to be exceptionally adaptable and resilient to ensure consistent security coverage. Real-world deployment should also assess how these models handle evolving security threats and whether they can effectively learn and adapt over time to counter new types of attacks. The robustness, scalability, and practical applicability of Distributed Reinforcement Learning models for IoT security can be thoroughly evaluated by conducting extensive testing in such variable and realistic conditions, providing valuable insights into their potential and limitations in real-world scenarios. Such comprehensive evaluations are vital for refining these models and establishing their readiness for widespread deployment in protecting IoT ecosystems.

6.9. Standardization and Interoperability

The diversity and vastness of IoT devices and networks present a considerable challenge in ensuring that Distributed Reinforcement Learning models work effectively across different systems. In the absence of common standards, the potential of these models to deliver comprehensive security solutions is significantly limited. Therefore, it is essential to establish universally accepted protocols and interfaces, enabling smooth communication and data exchange between IoT devices and networks. These standardized methods will not only facilitate the integration of Distributed Reinforcement Learning models into existing IoT infrastructures but will also enhance the scalability of these models. Distributed Reinforcement Learning models can be more readily adapted and scaled to suit different IoT environments, from small-scale home networks to large industrial systems, by ensuring compatibility and ease of integration.

Moreover, developing common frameworks and guidelines for Distributed Reinforcement Learning in IoT will contribute to security solutions' overall robustness and efficiency. With standardized protocols, Distributed

Reinforcement Learning models can more efficiently process data from diverse sources, leading to more accurate and timely security responses. This standardization also opens the door for collaborative development and sharing of best practices among different stakeholders, including IoT device manufacturers, network providers, and security experts. As a result, the collective knowledge and experience in IoT security can be leveraged, driving innovation and continuous improvement in Distributed Reinforcement Learning applications. Ultimately, the concerted effort to address standardization and interoperability challenges will not only enhance the deployment and effectiveness of Distributed Reinforcement Learning models in IoT security but will also pave the way for more advanced and secure IoT ecosystems.

6.10. Human-in-the-Loop Reinforcement Learning

Integrating Distributed Reinforcement Learning in IoT applications is not just about augmenting the data-driven models with human insights but about creating a synergistic relationship where human expertise and machine learning algorithms inform and enhance each other. Human operators and security experts bring a wealth of experience and intuition, understanding nuanced and emerging threats that might not yet be represented in data. By incorporating their feedback, Distributed Reinforcement Learning models can be trained to recognize these subtleties and respond to complex or novel threats. Furthermore, human feedback can guide the Distributed Reinforcement Learning models in prioritizing security issues, ensuring that the most critical vulnerabilities are addressed first. This collaboration can also lead to developing more robust and resilient IoT security solutions, as human experts can provide oversight and correct any biases or errors that the models might develop over time.

From the perspective of the practical implementation of human-in-the-loop systems in Distributed Reinforcement Learning models for IoT security, the key challenge lies in effectively integrating human feedback into the learning process without disrupting the model's ability to operate autonomously. This integration could take the form of periodic reviews where human experts evaluate the model's decisions and provide corrective feedback or a more dynamic interaction where human inputs are continuously fed into the model, allowing it to adapt in real-time. Such an approach would not only enhance the model's learning efficiency but also build trust in automated security systems. Developing interfaces and communication protocols that enable precise and effective interaction between human experts and Distributed Reinforcement Learning models is also essential. When executed effectively, this human-machine collaboration can greatly enhance the detection and mitigation of IoT security threats. This makes the systems more reliable and efficient in safeguarding against the continuously evolving landscape of cyber threats.

7. Conclusions

The rapid growth in the Internet of Things (IoT) has resulted in many interconnected devices across various networks. However, this expansion has also presented some serious security challenges, primarily because of IoT devices' inherent vulnerabilities and diverse nature. As a result, traditional security solutions cannot be readily deployed to address dynamic IoT threats.

Hence, this review article explores using Distributed Reinforcement Learning approaches to enhance IoT security in distributed and heterogeneous networks. Namely, the work explores the fundamental theories reinforcing IoT security, examines the key challenges facing Distributed Reinforcement Learning in this domain, and considers various potential solutions. Furthermore, the work also reviews the fundamentals of Distributed Reinforcement Learning, its advantages, disadvantages, and possible uses in tackling IoT security challenges.

The article also incorporates some case studies, experiments, and performance analysis to compare approaches based on Deep Reinforcement Learning with traditional methods. In conclusion, we have discussed future directions, emerging trends, and unresolved challenges in applying Distributed Reinforcement Learning for IoT security.

8. List of Abbreviations

A3C	Asynchronous Advantage Actor-Critic
AES	Advanced Encryption Standard
AI	Artificial Intelligence
DDoS	Distributed Denial-of-Service
DDPG	Deep Deterministic Policy Gradient
DoS	Denial-of-Service
DQN	Trained Deep Q-Network
E2E	End-to-End
GAN	Generative Adversarial Network
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
IDS	Intrusion Detection Systems
IoT	Internet of Things
LLMs	Large Language Models
MITM	Man-in-the-Middle
OTA	Over-the-Air
TLS	Transport Layer Security
VPNs	Virtual Private Networks

Author Contributions

Conceptualization, R.A., A.H., M.R., and S.K.J.; methodology, M.R., and S.K.J., and M.A.; software, M.R.; validation, M.R., S.K.J., and M.A.; formal analysis, S.K.J.; investigation, M.R.; resources, S.K.J.; data curation, M.R.; writing—original draft preparation, S.K.J., A.H., D.O., M.R.; writing—review and editing, M.R., H.D., D.O., M.A., A.H., and R.H.; visualization, S.K.J.; supervision, M.R.; project administration, M.R., D.O., and R.A. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials

Not applicable.

Consent for Publication

Not applicable.

Conflict of Interest

The authors declare no conflicts of interest.

Funding

This research received no external funding.

References

[1] I. H. Sarker, A. I. Khan, Y. B. Abushark, F. Alsolami, “Internet of things (IoT) security intelligence: A comprehensive overview, machine learning solutions and research directions,” *Mobile Networks and Applications*, vol. 28, pp. 296–312, 2022. [CrossRef]

[2] M. Rahouti, M. Ayyash, S. K. Jagatheesaperumal, D. Oliveira, “Incremental learning implementations and vision for cyber risk detection in iot,” *IEEE Internet of Things Magazine*, vol. 4, pp. 114–119, 2021. [CrossRef]

[3] S. S. Gill, S. Tuli, M. Xu, I. Singh, K. V. Singh, D. Lindsay, et al., “Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges,” *Internet of Things*, vol. 8, p. 100118, 2019. [CrossRef]

[4] S. Bagchi, T. F. Abdelzaher, R. Govindan, P. Shenoy, A. Atrey, P. Ghosh, et al., “New frontiers in iot: Networking, systems, reliability, and security challenges,” *IEEE Internet of Things Journal*, vol. 7, pp. 11330–11346, 2020. [CrossRef]

[5] A. Chehri, G. Jeon, F. Rivest, H. T. Mouftah, “Evolution and Trends in Artificial Intelligence of Things Security: When Good Enough is Not Good Enough! *IEEE Internet of Things Magazine*, vol. 5, pp. 62–66, 2022. [CrossRef]

[6] N. Quadar, A. Chehri, G. Jeon, M. M. Hassan, G. Fortino, “Cybersecurity Issues of IoT in Ambient Intelligence (AmI) Environment,” *Internet of Things Magazine*, vol. 5, pp. 140–145, 2022. [CrossRef]

[7] N. M. Karie, N. M. Sahri, P. Haskell-Dowland, “IoT threat detection advances, challenges and future directions,” In: Proceedings of the 2020 workshop on emerging technologies for security in IoT (ETSecIoT), Sydney, NSW, Australia, 21 April 2020, pp. 22–29.

[8] G. Jeon, A. Chehri, “Security Analysis Using Deep Learning in IoT and Intelligent Transport System,” in X. Qu, L. Zhen, R. J. Howlett, L. C. Jain (Eds.) *Smart Transportation Systems 2021*, Singapore: Springer, 2021; pp. 9–19.

[9] T. Salman, R. Jain, “Networking protocols and standards for internet of things,” in *Internet of Things and Data Analytics Handbook*, Hoboken: John Wiley & Sons, 2017, pp. 215–238.

[10] L. Atzori, A. Iera, G. Morabito, “Understanding the Internet of Things: Definition, potentials, and societal role of a fast evolving paradigm,” *Ad Hoc Networks*, vol. 56, pp. 122–140, 2017. [CrossRef]

[11] A. Das, P. Rad, K. K. R. Choo, B. Nouhi, J. Lish, J. Martel, “Distributed machine learning cloud teleophthalmology IoT for predicting AMD disease progression,” *Future Generation Computer Systems*, vol. 93, pp. 486–498, 2019. [CrossRef]

[12] F. Alwahedi, A. Aldhaheri, M. A. Ferrag, A. Battah, N. Tihanyi, “Machine learning techniques for IoT security: Current research and future vision with generative AI and large language models,” *Internet of Things and Cyber-Physical Systems*, vol. 4, pp. 167–185, 2024. [CrossRef]

[13] C. Ni, S. C. Li, “Machine learning enabled industrial iot security: Challenges, trends and solutions,” *Journal of Industrial Information Integration*, vol. 38, p. 100549, 2024. [CrossRef]

[14] A. Nuhu, A. F. M. Raffei, M. F. Ab Razak, A. Ahmad, “Distributed Denial of Service Attack Detection in IoT Networks using Deep Learning and Feature Fusion: A Review,” *Mesopotamian Journal of CyberSecurity*, vol. 4, pp. 47–70, 2024. [CrossRef]

[15] M. Le, T. Huynh-The, T. Do-Duy, T. H. Vu, W. J. Hwang, Q. V. Pham, “Applications of distributed machine learning for the Internet-of-Things: A comprehensive survey,” *arXiv* arXiv:2310.10549, 2023. [CrossRef]

[16] A. Uprety, D. B. Rawat, “Reinforcement learning for iot security: A comprehensive survey,” *IEEE Internet of Things Journal*, vol. 8, pp. 8693–8706, 2020. [CrossRef]

[17] W. Chen, X. Qiu, T. Cai, H. N. Dai, Z. Zheng, Y. Zhang, “Deep reinforcement learning for Internet of Things: A comprehensive survey,” *IEEE Communications Surveys & Tutorials*, vol. 23, pp. 1659–1692, 2021. [CrossRef]

[18] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, M. Guizani, “A survey of machine and deep learning methods for internet of things (IoT) security,” *IEEE Communications Surveys & Tutorials*, vol. 22, pp. 1646–1685, 2020. [CrossRef]

[19] M. S. Frikha, S. M. Gammar, A. Lahmadi, L. Andrey, “Reinforcement and deep reinforcement learning for wireless Internet of Things: A survey,” *Com-*

puter Communications, vol. 178, pp. 98–113, 2021. [CrossRef]

[20] S. Hu, X. Chen, W. Ni, E. Hossain, X. Wang, "Distributed machine learning for wireless communication networks: Techniques, architectures, and applications," *IEEE Communications Surveys & Tutorials*, vol. 23, pp. 1458–1493, 2021. [CrossRef]

[21] R. Ahmad, I. Alsmadi, "Machine learning approaches to IoT security: A systematic literature review," *Internet of Things*, vol. 14, p. 100365, 2021. [CrossRef]

[22] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, B. Sikdar, "A survey on IoT security: Application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019. [CrossRef]

[23] A. Chehri, I. Fofana, X. Yang, "Security Risk Modeling in Smart Grid Critical Infrastructures in the Era of Big Data and Artificial Intelligence," *Sustainability*, vol. 13, p. 3196, 2021. [CrossRef]

[24] P. Williams, I. K. Dutta, H. Daoud, M. Bayoumi, "A survey on security in Internet of Things with a focus on the impact of emerging technologies," *Internet of Things*, vol. 19, p. 100564, 2022. [CrossRef]

[25] P. K. Sadhu, V. P. Yanambaka, A. Abdelgawad, "Internet of Things: Security and Solutions Survey," *Sensors*, vol. 22, p. 7433, 2022. [CrossRef] [PubMed]

[26] H. Mrabet, S. Belguith, A. Alhomoud, A. Jemai, "A Survey of IoT Security Based on a Layered Architecture of Sensing and Data Analysis," *Sensors*, vol. 20, p. 3625, 2020. [CrossRef] [PubMed]

[27] Z. Cui, X. Fei, S. Zhang, X. Cai, Y. Cao, W. Zhang, J. Chen, "A hybrid blockchain-based identity authentication scheme for multi-WSN," *IEEE Transactions on Services Computing*, vol. 13, pp. 241–251, 2020. [CrossRef]

[28] M. A. Obaidat, S. Obeidat, J. Holst, A. Al Hayajneh, J. Brown, "A Comprehensive and Systematic Survey on the Internet of Things: Security and Privacy Challenges, Security Frameworks, Enabling Technologies, Threats, Vulnerabilities and Countermeasures," *Computers*, vol. 9, p. 44, 2020. [CrossRef]

[29] E. Fazeldehkordi, T. M. Grønli, "A Survey of Security Architectures for Edge Computing-Based IoT," *IoT*, vol. 3, pp. 332–365, 2022. [CrossRef]

[30] M. A. Khan, I. U. Din, T. Majali, B. S. Kim, "A Survey of Authentication in Internet of Things-Enabled Healthcare Systems," *Sensors*, vol. 22, p. 9089, 2022. [CrossRef]

[31] A. Falayi, Q. Wang, W. Liao, W. Yu, "Survey of Distributed and Decentralized IoT Securities: Approaches Using Deep Learning and Blockchain Technology," *Future Internet*, vol. 15, p. 178, 2023. [CrossRef]

[32] K. Ragothaman, Y. Wang, B. Rimal, M. Lawrence, "Access Control for IoT: A Survey of Existing Research, Dynamic Policies and Future Directions," *Sensors*, vol. 23, p. 1805, 2023. [CrossRef]

[33] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf, Y. A. Bangash, "An In-Depth Analysis of IoT Security Requirements, Challenges, and Their Countermeasures via Software-Defined Security," *IEEE Internet of Things Journal*, vol. 7, pp. 10250–10276, 2020. [CrossRef]

[34] I. Lodha, L. Kolur, K. S. Hari, P. Honnavalli, "Secure wireless internet of things communication using virtual private networks," in *International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2019*, Springer: Singapore, 2020, pp. 735–742.

[35] M. Husnain, K. Hayat, E. Cambiaso, U. U. Fayyaz, M. Mongelli, H. Akram, et al., "Preventing mqtt vulnerabilities using iot-enabled intrusion detection system," *Sensors*, vol. 22, p. 567, 2022. [CrossRef] [PubMed]

[36] M. Rana, Q. Mamun, R. Islam, "Lightweight cryptography in IoT networks: A survey," *Future Generation Computer Systems*, vol. 129, pp. 77–89, 2022. [CrossRef]

[37] R. C. J. Neto, P. Mérindol, F. Theoleyre, "Enabling privacy by anonymization in the collection of similar data in multi-domain IoT," *Computer Communications*, vol. 203, pp. 60–76, 2023. [CrossRef]

[38] N. M. Karie, N. M. Sahri, W. Yang, C. Valli, V. R. Kebande, "A Review of Security Standards and Frameworks for IoT-Based Smart Environments," *IEEE Access*, vol. 9, pp. 121975–121995, 2021. [CrossRef]

[39] K. Y. Lam, S. Mitra, F. Gondesen, X. Yi, ANT-Centric IoT Security Reference Architecture—Security-by-Design for SatelliteEnabled Smart Cities," *IEEE Internet of Things Journal*, vol. 9, pp. 5895–5908, 2022. [CrossRef]

[40] F. Chen, D. Luo, J. Li, V. C. M. Leung, S. Li, J. Fan, "Arm PSA-Certified IoT Chip Security: A Case Study," *Tsinghua Science and Technology*, vol. 28, pp. 244–257, 2023. [CrossRef]

[41] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, A. Zanella, "IoT: Internet of Threats? A Survey of Practical Security Vulnerabilities in Real IoT Devices," *IEEE Internet of Things Journal*, vol. 6, pp. 8182–8201, 2019. [CrossRef]

[42] S. Shen, T. Zhu, D. Wu, W. Wang, W. Zhou, "From distributed machine learning to federated learning: In the view of data privacy and security," *Concurrency and Computation: Practice and Experience*, vol. 34, p. e6002, 2022. [CrossRef]

[43] K. Arulkumaran, M. P. Deisenroth, M. Brundage, A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, pp. 26–38, 2017. [CrossRef]

[44] I. Sharafaldin, A. H. Lashkari, A. A. Ghorbani, V. Michel, B. Thirion, O. Grisel, et al., "Toward generating a new intrusion detection dataset and intrusion traffic characterization," *ICISSp*, vol. 1, pp. 108–116, 2018.

[45] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, A. A. Ghorbani, CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT

environment," *Sensors*, vol. 23, p. 5941, 2023. [CrossRef] [PubMed]

[46] E. Owusu, M. Rahouti, S. K. Jagatheesaperumal, K. Xiong, Y. Xin, L. Lu, et al., "Online Network DoS/DDoS Detection: Sampling, Change Point Detection, and Machine Learning Methods," *IEEE Communications Surveys & Tutorials*, 2024. [CrossRef]

[47] M. Tavallaee, E. Bagheri, W. Lu, A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set". in: Proceedings of the 2009 IEEE symposium on computational intelligence for security and defense applications, Ottawa, ON, Canada, 8–10 July 2009, pp. 1–6.

[48] N. Moustafa, J. Slay, UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). in: Proceedings of the 2015 military communications and information systems conference (MilCIS), Canberra, Australia, 10–12 November 2015, pp. 1–6.

[49] S. Garcia, A. Parmisano, M. J. Erquiaga, *IoT-23: A labeled dataset with malicious and benign IoT network traffic*, Technic Report, Praha: Stratosphere Lab., 2020.

[50] S. Wang, M. Chen, Z. Yang, C. Yin, W. Saad, S. Cui, et al., "Distributed reinforcement learning for age of information minimization in real-time IoT systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, pp. 501–515, 2022. [CrossRef]

[51] Z. Li, C. Huang, S. Deng, W. Qiu, X. Gao, "A soft actor-critic reinforcement learning algorithm for network intrusion detection," *Computers & Security*, vol. 135, p. 103502, 2023. [CrossRef]

[52] S. Bahrami, Y. C. Chen, V. W. Wong, "Deep reinforcement learning for demand response in distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, pp. 1496–1506, 2020. [CrossRef]

[53] A. Kumar, D. Singh, "Detection and prevention of DDoS attacks on edge computing of IoT devices through reinforcement learning," *International Journal of Information Technology*, vol. 16, pp. 1365–1376, 2024. [CrossRef]

[54] S. Deng, Z. Xiang, P. Zhao, J. Taheri, H. Gao, J. Yin, et al., "Dynamical resource allocation in edge for trustable internet-of-things systems: A reinforcement learning method," *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 6103–6113, 2020. [CrossRef]

[55] S. Mishra, S. Chakraborty, K. S. Sahoo, M. Bilal, "Cogni-Sec: A secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system," *Internet of Things*, vol. 24, p. 100978, 2023. [CrossRef]

[56] A. Ferdowsi, W. Saad, "Generative adversarial networks for distributed intrusion detection in the internet of things. in: Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019, pp. 1–6.

[57] H. Benaddi, M. Jouhari, K. Ibrahimi, J. Ben Othman, E. M. Amhoud, "Anomaly detection in industrial IoT using distributional reinforcement learning and generative adversarial networks," *Sensors*, vol. 22, p. 8085, 2022. [CrossRef]

[58] T. Mohammed, A. Albeshri, I. Katib, R. Mehmood, "UbiPriSEQ—Deep reinforcement learning to manage privacy, security, energy, and QoS in 5G IoT hetnets," *Applied Sciences*, vol. 10, p. 7120, 2020. [CrossRef]

[59] F. Louati, F. B. Ktata, I. Amous, "Big-IDS: A decentralized multi agent reinforcement learning approach for distributed intrusion detection in big data networks," *Cluster Computing*, vol. 27, pp. 6823–6841, 2024. [CrossRef]

[60] Z. Abou El Houda, H. Moudoud, B. Brik, "Federated Deep Reinforcement Learning for Efficient Jamming Attack Mitigation in O-RAN," *IEEE Transactions on Vehicular Technology*, vol. 73, pp. 9334–9343, 2024. [CrossRef]

[61] A. A. Diro, N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, 2018. [CrossRef]

[62] A. N. Bikos, S. Kumar, "Reinforcement learning-based anomaly detection for Internet of Things distributed ledger technology. in: Proceedings of the 2021 IEEE Symposium on Computers and Communications (ISCC), Athens, Greece, 5–8 September 2021, pp. 1–7.

[63] M. Goudarzi, M. Palaniswami, R. Buyya, "A distributed deep reinforcement learning technique for application placement in edge and fog computing environments," *IEEE Transactions on Mobile Computing*, vol. 22, pp. 2491–2505, 2021. [CrossRef]

[64] X. Feng, J. Han, R. Zhang, S. Xu, H. Xia, "Security defense strategy algorithm for Internet of Things based on deep reinforcement learning," *High-Confidence Computing*, vol. 4, p. 100167, 2024. [CrossRef]

[65] V. Brindha Devi, N. M. Ranjan, H. Sharma, "IoT attack detection and mitigation with optimized deep learning techniques," *Cybernetics and Systems*, vol. 55, pp. 1702–1728, 2024. [CrossRef]

[66] X. Liu, W. Yu, F. Liang, D. Griffith, N. Golmie, "On deep reinforcement learning security for Industrial Internet of Things," *Computer Communications*, vol. 168, pp. 20–32, 2021. [CrossRef]

[67] G. D. L. T. Parra, P. Rad, K. K. R. Choo, N. Beebe, "Detecting Internet of Things attacks using distributed deep learning," *Journal of Network and Computer Applications*, vol. 163, p. 102662, 2020. [CrossRef]

[68] G. Rjoub, O. A. Wahab, J. Bentahar, A. Bataineh, "Trust-driven reinforcement selection strategy for federated learning on IoT devices," *Computing*, vol. 106, pp. 1273–1295, 2024. [CrossRef]

[69] Q. Miao, H. Lin, X. Wang, M. M. Hassan, "Federated deep reinforcement learning based secure data sharing for Internet of Things," *Computer Networks*, vol. 197, p. 108327, 2021. [CrossRef]