energies

MDPI

Review

A Review of Denial of Service Attack and Mitigation in the Smart Grid Using Reinforcement Learning

Ines Ortega-Fernandez [1,2,3] and Francesco Liberati [4,*]

[1] Galician Research and Development Center in Advanced Telecommunications (GRADIANT), 36310 Vigo, Spain
[2] CITMAga, 15782 Santiago de Compostela, Spain
[3] Escola de Enxeñaría de Telecomunicación, Universidade de Vigo, 36310 Vigo, Spain
[4] Department of Computer Control and Management Engineering (DIAG) "Antonio Ruberti", University of Rome "La Sapienza", Via Ariosto, 25, 00185 Rome, Italy
* Correspondence: liberati@diag.uniroma1.it

Abstract: The smart grid merges cyber-physical systems (CPS) infrastructure with information and communication technologies (ICT) to ensure efficient power generation, smart energy distribution in real-time, and optimisation, and it is rapidly becoming the current standard for energy generation and distribution. However, the use of ICT has increased the attack surface against the electricity grid, which is vulnerable to a wider range of cyberattacks. In particular, Denial-of-Service (DoS) attacks might impact both the communication network and the cyber-physical layer. DoS attacks have become critical threats against the smart grid due to their ability to impact the normal operation of legitimate smart-grid devices and their ability to target different smart grid systems and applications. This paper presents a comprehensive and methodical discussion of DoS attacks in the smart grid, analysing the most common attack vectors and their effect on the smart grid. The paper also presents a survey of detection and mitigation techniques against DoS attacks in the smart grid using reinforcement learning (RL) algorithms, analysing the strengths and limitations of the current approaches and identifying the prospects for future research.

Keywords: smart grids; cyberattacks; denial-of-Service; reinforcement learning; cyber detection

check for updates

Citation: Ortega-Fernandez, I.; Liberati, F. A Review of Denial of Service Attack and Mitigation in the Smart Grid Using Reinforcement Learning. *Energies* **2023**, *16*, 635. https://doi.org/10.3390/en16020635

Academic Editors: Yanbin Qu and Huihui Song

Received: 17 November 2022
Revised: 29 December 2022
Accepted: 2 January 2023
Published: 5 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The smart grid is the current energy management and distribution trend: it merges cyber–physical systems (CPS) infrastructure with information and communication technologies (ICT) to ensure efficient power generation, smart energy distribution in real-time, and optimisation [1]. It also allows for greater integration of alternative energy sources such as solar and wind power, which are heavily reliant on weather patterns. Smart grid applications include extraction of business value, smart charging of electric vehicles, smart distribution, generation and storage of energy, grid optimization, grid self-healing with fault protection technology, and many others [2] (Figure 1). However, the use of ICT introduces new threats to the smart grid infrastructure and makes it vulnerable to cyber-attacks: using legacy technologies such as conventional Supervisory Control and Data Acquisition (SCADA) systems or running most CPS protocols over TCP/IP exposes the smart grid to attack vectors found in traditional information systems [3].

Denial-of-Service (DoS) attacks, in particular, have become critical threats to the smart grid because they target the availability of the grid infrastructure and services: in the context of smart grids, this includes both "ensuring timely and reliable access to and use of information" [4] and "ensuring access to enough power" [5]. Since the network lacks extensive storage capacity, the generated electrical power must be consumed in a short period of time. A DoS attack could prevent grid measurements from reaching the control centre, so affecting the frequency equilibrium between power generation and

consumption [6]: the control centre uses data gathered from multiple sections of the smart grid to determine energy requirements, provide data to energy providers for billing, and for controlling consumption and generation of electricity. Furthermore, any disruption in the network must be addressed quickly in order to avoid major service interruptions. As a result, advanced defence mechanisms that address the special constraints of real-time operation and availability of the smart grid are required to protect against DoS attacks.

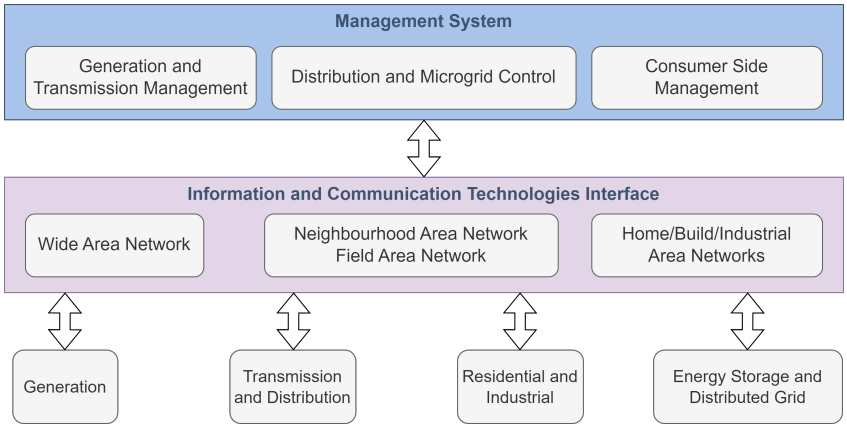

Figure 1. Smart Grid infrastructure and components.

Huseinovi et al. provide in [5] a taxonomy of the major power grid applications subject to DoS attacks:

- The Advanced Metering Infrastructure provides smart meters with bidirectional communication capabilities and data transfer with the control centre. It is a common target of DoS attacks.
- The Distribution Management System monitors, protects, controls and optimizes the assets of the distribution grid, and might be affected by load frequency disturbance caused by a DoS attack.
- Wide Area Monitoring, Protection and Control Systems are also subject to DoS attacks [7,8].
- Demand Side Management might be affected by DoS attacks that target the devices in charge of maintaining the load and supply balance from the demand side.
- At last, the Energy Management System is in charge of keeping the balance between the energy supply and the demand. A Distributed DoS (DDoS) targeting the Energy Management System will prevent it from controlling the power ratio between consumption and generation, causing problems such as voltage drop/rise.

Due to the large diversity of available cyber–physical layer protocols, the use of an open communication network geographically distributed, and the limited computational abilities of the smart grid devices, among others, securing the smart grid is still an open challenge [9]. While basic security measures (such as authentication mechanisms, encrypted communications or the use of firewalls) can be effective to address simple attacks, advanced threats require continuous monitoring, detection and prevention systems [9], and a quick and efficient response to incidents. In particular, Intrusion Prevention Systems (IPS) can be employed to detect and mitigate DoS attacks by executing automated mitigation actions when a cyberattack is detected, for example, by re-configuring firewall rules, the network topology (in the case of Software Defined Networks (SDNs)) or by implementing different actions in the control layer [10].

Traffic monitoring tools might be used to obtain statistical information about the data exchanged in the ICT interface of the smart grid in order to detect cyberattacks. In particular, the study of traffic flows might be useful to detect DoS attacks that cause packet delays and communication network congestion, and even to detect the presence of new devices in the network. IPSs might be classified into two big groups (Figure 2):

1. Signature-based IPSs, which use signatures and patterns of well-known DoS attacks to compare current network traffic with its expected pattern, raising an alarm when the current behaviour does not match the learned signature or rule. Although these methods are easy to implement, they fail to detect novel or unknown attacks [11].
2. Anomaly-based IPSs, which learn a pattern of the normal behaviour of a network by means of statistical properties, and raise an alarm when the current behaviour does not match the expected pattern, allowing the IPS to detect unknown attacks [12]. However, anomaly-based IPSs are more costly to train and tune, and it is more difficult to obtain the exact root cause of the detected anomaly. The pattern of legitimate behaviour may be learned with a variety of techniques: traditionally, pattern-based intrusion detection has been performed by analysing the contents of each individual network packet to find anomalies that deviate from the learned pattern, using a set of techniques named Deep Packet Inspection. However, inspecting each packet is not efficient in large networks, and is even impossible at network speeds of Gigabits per second. The main alternative to Deep Packet Inspection is flow-based anomaly detection, where the communication patterns (in Netflow [13] or IPFIX [14] format) are studied, instead of the content of each individual packet [15].

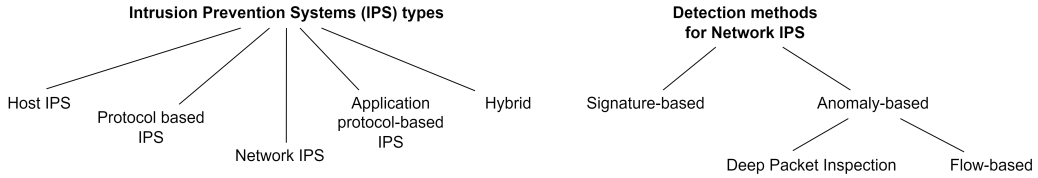

Figure 2. Network Intrusion Prevention Systems types and detection methods.

This work will focus on two key topics: first, an examination of the most frequent DoS attack vectors and their effect on the smart grid, addressing the most relevant types of DoS attacks targeting the Smart Grid, namely [5]:

- **Flooding attacks**, which overwhelm the communication network with packets to disturb legitimate communication.
- **Jamming attacks**, which interfere with the wireless signals at the physical layer to deny or delay the communication between smart grid devices.
- **De-synchronization attacks**, which target smart grid systems which rely on exact timing and synchronized measurements.
- **Amplification attacks**, which take advantage of networking protocols to overwhelm the communication network or exhaust device resources.
- **False Data Injection attacks**, which alter the packet content with the aim of disrupting smart grid services.

Second, the most promising RL-based techniques for detecting and mitigating the aforementioned cyberattacks are discussed, identifying the gaps in the literature and prospects for future research. The main requirement to implement IPSs in the smart grid is to incorporate scalable, resilient and efficient algorithms that do not interrupt normal smart grid operations and are able to handle a variety of network devices, topology and infrastructures. RL algorithms have the ability to obtain optimal action policies, learning how to map observations taken from the environment to actions to maximize an expected future reward [16]. IPSs based on RL are a current trend for securing complex systems which require autonomous agents capable of learning to make decisions. RL-based IPS are good candidates to detect and mitigate different DoS attacks in the Smart Grid given the flexibility provided by RL algorithms to learn the best course of action against various situations and attacks.

The remainder of this paper is structured as follows: Section 2 describes the different types of DoS attacks targeting the smart grid, classifying them in terms of the applications they target, the used protocols, and the main defence mechanisms. Section 3 addresses the

most novel defence and mitigation techniques using RL. Finally, Section 4 summarises the main conclusions gathered from the survey, identifying the prospects for future research.

2. DoS Attacks in the Smart Grid

The smart grid is made up of various elements that form a hierarchical architecture: in general, a set of measurement components (such as smart meters or Programmable Logic Controllers) gather data from the environment and send it to a control centre via communication protocols that run over TCP/IP, inheriting the DoS attack vectors from the internet domain. Because the control centre and the physical layer are usually geographically separated, the cyber–physical infrastructure shares the networking infrastructure with the Internet, allowing attackers and cyber-criminals to gain access to the smart grid. DoS attacks are one of the attacks that can have a greater impact on the smart grid: a DoS attack against the smart grid has the primary effect of disrupting or delaying power delivery; the attack could target a single device or the various sections of the smart grid: generation, transmission, distribution, and consumption [5].

DoS attacks against communication between smart metering equipment and control centres, in particular, may stop signals from reaching their destinations on time, preventing the control centre from maintaining a strong situational awareness of the grid's condition, and thus leading to grid instability. The attackers may employ many devices (and botnets) to carry out what is called a Distributed DoS attack, and use spoofed IP addresses to mask their identities. This section summarises the most relevant DoS active attack types against the smart grid, which might cause grid instability and/or unavailability. The attacks are classified based on the technique used, and the consequences for the smart grid are discussed. Table 1 summarises the main attack techniques, their targets in the smart grid infrastructure and the main mitigation and detection techniques. This study does not cover passive attacks, where the attacker eavesdrops on communications or analyses traffic in order to obtain information about the smart grid, but without modifying data or interacting with the infrastructure.

Table 1. Summary of the main DoS attacks against the smart grid, relevant works and main identified defence strategies.

Attack	Target	Defence Strategy	Relevant Works
Flooding	ICMP UDP TCP	Network monitoring with NIDS Moving Target Defence Drop or filter traffic Anomaly Detection	[17] [18] [19] [20] [21] [22]
Jamming	Wireless Communication Layer	Signal Strength Measurements Monitoring of the Carrier Sensing Time Threshold based detection on the PDR Consistency Checks Network monitoring with NIDS Delayed Disconnect Intelligent selection of wireless channels Drop/filter traffic	[23] [24] [25] [26] [26]
De-synchronization	Global Positioning System (GPS) Network Time Protocol (NTP)	Monitor system stability Monitor the GPS carrier-to-noise ratio Use IEEE 1588-2008 precision time protocol Use stable atomic clocks	[27] [28] [29]
Amplification	UDP	Network monitoring with NIDS Filter or drop traffic Deep Packet Inspection Anomaly detection	[30]
False Data Injection	Unencrypted, unauthenticated communications	Deep Packet Inspection Anomaly Detection	[31] [32] [33]

2.1. Flooding Attacks

A flooding attack tries to overwhelm the network, resulting in delayed or interrupted communication between legitimate devices. It might be performed by taking advantage of multiple network protocols, such as ICMP, UDP, or TCP. The attack is simple to execute but causes significant network disruption.

TCP SYN Flooding [34] takes advantage of the three-way handshake mechanism of the TCP/IP protocol, and therefore all smart grid protocols running on top of TCP/IP (such as Modbus TCP) are vulnerable to SYN flood attacks. The attacker sends an SYN request packet to the victim, which replies with an SYN/ACK packet and keeps a port open waiting for an ACK packet from the attacker to establish the connection. However, the attacker never sends the ACK packet, forcing the client to keep the port open until the connection expires (Figure 3). The main outcome of a successful SYN flooding attack is that the victim runs out of available ports to initiate new connections with legitimate devices.

UDP Flooding attacks are executed when an attacker generates a high amount of packets to random destination ports of the victim. If the port is closed, the victim will respond with an ICMP packet; if the amount of generated ICMP packets is large enough, it might overwhelm the network preventing legitimate packets from reaching their destination: in [17], Asri et al. show how a UDP flooding attack could take down the entire grid infrastructure. The UDP attack prevents the control centre from gathering usage data from the grid, and as a result, the power plant stops producing electricity, causing the entire network to fail. Likewise, Ping flooding [35] tries to overwhelm the network with ICMP packets: the targeted device becomes overloaded with ICMP Echo Request (*ping*) packets. The attacker tries to consume all the available bandwidth, preventing legitimate packets from reaching their destination.

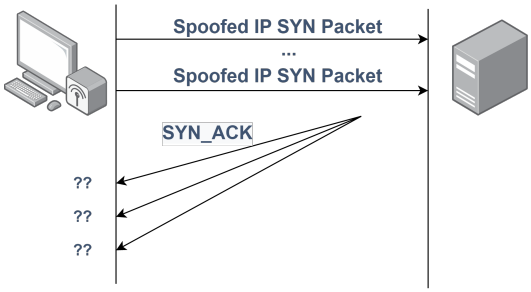

Figure 3. Conceptual diagram of the TCP SYN Flood attack. The attacker uses spoofed IP addresses, so the victim never gets a response to the SYN-ACK packets, forcing it to keep open ports and exhausting its resources.

Depending on the role of the victim in the smart grid, the attack might have different consequences in the physical environment. Flooding attacks might increase delays on time-critical messages, such as those exchanged between the control centre and smart meters [19–21]. The DoS attack might not only interrupt communications but also exhaust the victim's resources in terms of CPU consumption, preventing it from performing legitimate tasks [22]. Common defence strategies include the deployment of network monitoring through Network Intrusion Detection Systems (NIDS) and anomaly detection [17,19,21], or the use of advanced moving target defence mechanisms as proposed in [18].

2.2. Jamming Attacks

A recent DoS attack vector in wireless networks is initiated by jamming the signals at the physical layer to deny or delay the communication between smart grid devices [36]. The attackers might use a variety of techniques, from the simple continual transmission of interference signals to advanced attacks that exploit vulnerabilities on application layer protocols [23]. In [24], Temple et al. implement a jamming attack with two different goals: to deny the electrical service during a certain time window, and to produce physical disturbance of power grid frequency by causing *load shedding*, assuming an attacker with perfect knowledge of the infrastructure. Li et al. investigate in [25] both jamming and anti-jamming techniques in a multichannel wireless network that connects remote sensors and the control centre in a smart grid, modelling the interaction between the grid sensors and the attacker as a zero-sum stochastic game.

The detection and mitigation techniques for jamming attacks usually study parameters associated with the stability of the signals in the network, such as the signal strength indicator and the packet loss rate [26]. Other strategies are based on consistency checks, delayed disconnect, or detection of media access control layer misbehaviour through network monitoring [23], the intelligent selection of the communication channel to avoid the use of channels that are under attack [25].

2.3. De-Synchronization Attacks

Since many smart grid applications depend on synchronous measurements, they rely on exact timing information. Spoofing the Global Positioning System (GPS) signals is one method of carrying out de-synchronization attacks in the smart grid. Most measuring equipment employs the GPS to obtain exact timing: characteristics such as frequency and voltage are often sampled on a regular basis thanks to the GPS timing signal, and the measurements are aligned to a common time domain by the control centre. A de-synchronization attack occurs when a malicious attacker alters the sampling time by forging a GPS signal, causing the measuring device to sample the signal at the incorrect time (Figure 4). The misaligned measurement reaches the control centre, which acquires an inaccurate grid status. Zang et al. investigate in [27] the smart grid de-synchronization threats in three smart grid applications: transmission line fault detection and location, voltage stability monitoring, and event location. They demonstrate how a time de-synchronization attack might degrade such applications' performance, resulting in erroneous operations in the smart grid.

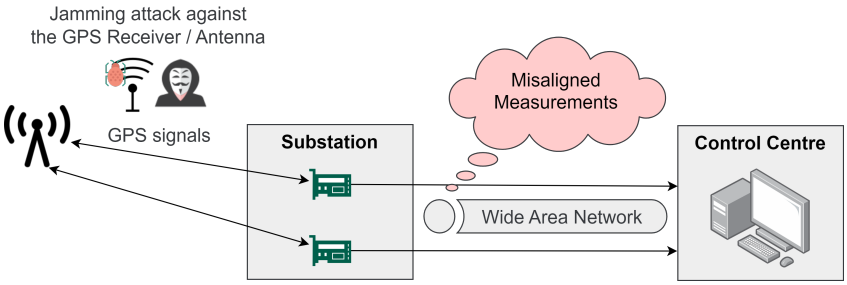

Figure 4. Example of a jamming attack on the GPS signal, which causes misaligned measurements to arrive at the control centre.

A different set of de-synchronization attacks in the smart grid are performed by exploiting vulnerabilities in the Precision Time Protocol (PTP). PTP is one of the IEEE 1588 protocols that permit time synchronization between devices with varied clock resolutions, precision, and stability with microsecond accuracy [37]. PTP is widely used in the smart grid at the substation level to obtain sample values with 1 µs accuracy. The PTP master is connected to the substation bus: the master receives timing from a GPS signal and distributes accurate time reference to all connected devices via synchronization messages under PTP [29]. However, PTP is vulnerable to different attacks, including DoS, packet manipulation and selective packet delay [28]. The authors of [29] exploit the PTP protocol, by introducing a variable delay in the PTP master communication path, and manipulating the clock of the connected devices. A delay attack against PTP will manipulate the clocks of all connected devices, affecting the functionality of merging units and potentially targeting all applications relying on precise timing, such as sampled values, fault localization, differential protection, or synchrophasor measurements.

The defence and mitigation techniques against jamming attacks include the use of cross-layer monitoring of the GPS carrier-to-noise ratio to detect time de-synchronization attacks [38], or the use of highly stable atomic clocks or time synchronization with the precision time protocol defined by IEEE 1588-2008 [39]. Moreover, authentication mechanisms can be used to prevent spoofing, which is the main enabler of de-synchronization attacks.

2.4. Amplification Attacks

Amplification attacks are a kind of volumetric DoS attack which involves reflection and amplification: the attacker spoofs an IP address (reflection), while exploiting UDP-based protocols that provide a much larger response than the request from the attacker (amplification) to overwhelm a network [40]. In contrast to flooding attacks, amplification attacks consume fewer resources from the attacker side but are more difficult to implement. Amplification attacks have four main characteristics [41]:

- Distributed: usually, multiple servers using the UDP protocol are used to launch the attack.
- Camouflage: attackers spoof their IP addresses into the addresses of the victim. Victims receive a lot of traffic from amplifiers (the server that is abused by the attackers).
- Reflexivity: the traffic is never received directly from the attacker, but indirectly by the amplifier's reflection.
- Amplification: the traffic reflected from the amplifier servers to the victims is much larger than the traffic sent to amplifiers from the attackers.

The main three types of amplification attacks targeting the smart grid are DNS, NTP and SNMP amplification. The Domain Name System (DNS) protocol translates domain names into IP addresses. In the context of smart grids, any device connected to a supervisory, control and data acquisition system has an IP address that is stored in the DNS server. In a DNS amplification attack, the attacker sends UDP packets with forged IP addresses to a DNS resolver, which acts as an amplifier server. The forged IP is the victim's IP. Each UDP packet requests the DNS resolver to send the largest response possible (by sending the "ANY" argument). When the DNS resolver receives the request, it sends to the victim a large response, overwhelming the networks and causing service interruption [42].

Likewise, a Network Time Protocol (NTP) amplification attack uses an NTP server as an amplifier: in the smart grid, NTP servers are used to perform time synchronization between current and voltage measurements from different devices in the grid. In the NTP amplification case, the attacker creates a reflection attack between the master nodes (that receive packets) and the slave nodes in the substations. The attacker sends UDP packets with forged IP addresses to the NTP server using the "monlist" command, which forces the server into responding with the latest 600 IP addresses that have made requests to the NTP server. The IP in the UDP packets is, again, the victim's IP, which receives a large UDP packet overwhelming the network. Finally, the Simple Network Management Protocol (SNMP) is widely used in management consoles dedicated to manage and maintain Programmable Logic Controllers. If an attacker gains access to an SNMP server it would be able to use it to scan the network and create a list of local devices, which will become the amplifiers of the attack. The attacker forges UDP packets, requesting the devices to respond with as much data as possible: the SNPM server (in this case the victim) will receive a large volume of data from all the devices, becoming overwhelmed by the amount of petitions and data [43].

In [30], Yang et al. propose an intrusion detection system specific for synchrophasor measurements, capable of detecting man-in-the-middle and amplification attacks by combining protocol-based whitelists with behavioural anomaly detection, by performing deep packet inspection on the network frames. The consequences (and the defence and mitigation techniques) of amplification attacks are similar to those of flooding attacks: the main objective of amplification attacks is to saturate the available bandwidth of a network (or the processing capabilities of the device processing the network packets) with numerous and large network packets, targeting the network layer of the TCP/IP protocol stack.

2.5. False Data Injection Attacks

In this type of attack, the malicious actor intercepts the communication network traffic (for example, by sniffing unencrypted communications from the network) and extracts the actual values of the network frame. The attacker forges false packets to force the control centre into executing wrong actions of various types. While False Data Injection

(FDI) attacks target the integrity of the network packets, they might also be used as a DDoS tool [31,44] when they cause interruptions on different smart grid applications. A common assumption is that FDI attacks require complete knowledge of the grid topology. However, recent works show how an attacker with limited knowledge is also able to successfully perform FDI attacks [45]. In [31], Vuković et al. consider an attacker that manipulates the data exchanged between the control centre and the neighbouring nodes. The attacker successfully manipulates the data to disable the distributed state estimation system, preventing it from finding correct state estimates. They show the impact of the attack on the IEEE 118 bus power system, where the FDI attack prevents the distributed state estimation from converging, leading to DoS due to erroneous state and power flow estimations, preventing the control centre from taking adequate actions.

The most critical estimations, and therefore the main targets of FDI attacks in a smart grid, are energy demand, energy supply, grid-network states and electricity pricing estimation. FDI attacks result in abnormal state estimations and might be detected by performing Deep Packet Inspection, anomaly detection, or through system-theoretic approaches [32]. In addition, when the FDI attack targets distributed state estimation (as discussed in [31]), a distributed detection approach is recommendable: a mitigation strategy is presented in [31], based on fully distributed attack detection (which is able to understand which region of the grid is impacted by the FDI attack), followed by a mitigation algorithm that isolates the attacked region, so that the distributed state estimation can converge. In [33], Zang et al. propose to detect FDI attacks using deep autoencoders and generative adversarial networks to learn the unconformity between abnormal and normal measurements; they use deep autoencoders to reduce the dimensionality of the input data, which serves as the input of the generative adversarial network for anomaly detection.

3. Detection and Mitigation Mechanisms against Cyberattacks with RL

In recent years, RL approaches have gained popularity due to their ability to obtain optimal action policies in different domains. RL is a branch of artificial intelligence which focus on "what to do", i.e., on how to map situations to actions to maximize an expected reward. The algorithm is not told which actions to execute, but it must discover which actions are better (yield the most reward) in each situation by interacting with the environment. RL algorithms might be implemented with a variety of techniques (summarised on Figure 5), from Markov decision processes to deep neural networks. In particular, the combination of deep neural networks and RL (deep RL) provides great advantages: deep neural networks are able to approximate the values of the optimal states and action pairs and can handle bigger action/state spaces than traditional RL approaches.

The common use case for RL in the smart grid is to determine an optimal strategy to manage different aspects of the cyber–physical layer. For example, implementing a deep RL algorithm to learn the optimal strategy for real-time electric vehicle charging scheduling [46,47], or to use of a model-free RL method for load frequency control resistant to weather uncertainties [48]. Another studied use case is the use of RL approaches to model the presence of an attacker in a smart grid. In this setting, the RL agent represents the attacker, which uses RL to find the optimal attack vector against the grid [49,50].

However, there is limited work proposing RL approaches to find the best defence strategy in a smart grid against a cyberattack. This section reviews and discusses current approaches to DoS attack detection and/or mitigation using RL. We will focus on the strengths and the limitations of each proposed technique (that we summarise in Table 2), also identifying the prospects for future research.

Feng et al. model, in [51], the physical state, control inputs, and disturbances of a CPS. The defence scheme is set up as a zero-sum game where the defender is an actor-critic deep RL algorithm (a game-theoretical actor-critic neural network) that learns an optimal strategy to timely defend a CPS from unknown attacks. However, their scenario only considers the physical layer and does not evaluate the impact of the attack on the communication layer.

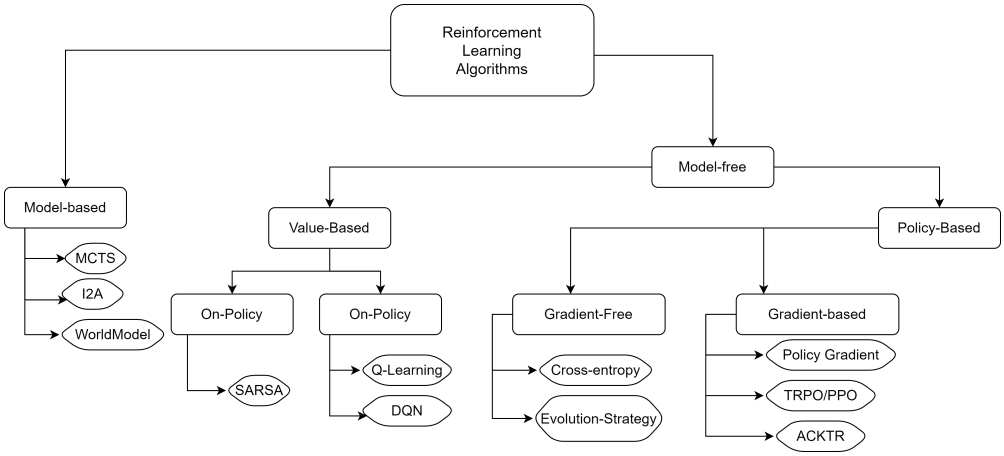

Figure 5. Taxonomy of popular RL algorithms.

In [52] a deep Q network detection scheme is proposed to defend against data integrity attacks in alternate current power grids. The objective of the attacker is to interrupt the normal operation of the power grid, bypassing the implemented bad data detection mechanisms. They use a deep Q network detection scheme where the state space of the smart grid is inferred by the agent (under attack or normal operation), and the action space includes that the system should be kept running (because the state observation did not detect any attack), or the system should be stopped. However, they assume an attack model where an attacker has some previous knowledge of the power grid system to be able to generate malicious data. Their deep Q network detection scheme improves the results of Feng et al. [51], in terms of delay-alarm error rates and detection failure rate, in three different IEEE bus systems, in both continuous and discontinuous attack models. However, the time complexity of the proposed approach is considerably high and grows exponentially with the number of devices in the power system, affecting the practicability of the prevention strategy in a real scenario.

Liu et al. implement in [53] a deep RL mitigation framework, specifically designed for DDoS flooding attacks against SDNs. They follow the Deep Deterministic Policy Gradient (DDPG) paradigm, where a parameterized actor function is maintained to define the current policy, mapping states to specific actions deterministically. They monitor the communication network load, and their action space is based on applying throttling to reduce the impact of the DDoS attack, dropping excess traffic. The agent tries to maximize the available bandwidth in the SDN by using a reward function on the ratio of benign and attacker traffic reaching the central SDN server. Therefore, the three goals of the agent are to prevent the server from crashing, allow benign traffic to reach the server, and prevent as much malicious traffic as possible. Even though this work is not focused on the smart grid, SDN is being used nowadays in smart grids to provide a resilient and flexible communication architecture.

Kurt et al. [54] formulate an attack/anomaly detection problem with a partially observable Markov decision process (POMDP), with a model-free RL algorithm that is able to defend from cyberattacks in an online manner without previous knowledge of any attack model. The defender learns low-magnitude attack models in order to become sensitive to even a very slight deviation from normal measurements. In order to detect an impending attack, the defender must first identify when an attack has begun. Because the agent cannot know when the attacker will initiate an attack, it must consider two states: a pre-attack and a post-attack state. The defender can take two actions after observing a measurement: stop and declare an attack, or continue and obtain additional measurements. The goal is for the agent to lessen its detection delay and false alarm rate by choosing actions that lead to high rewards. If it is in a pre-attack state and takes action to stop, it will receive a reward equal to one unit; however, if it takes action to continue obtaining measurements while already in a post-attack state, it will be penalized by receiving a cost equal to the detection delay.

The work of Wei et al. [55] presents a cyber-attack recovery (mitigation) strategy for the smart grid based on DDPG to optimally reclose the transmission lines when an attack is successful. In this case, the attackers can trip the smart grid circuit breakers in three different ways using FDI attacks: by modifying the GOOSE; redirecting local sampled measured values; or crafting false manufacturing message specification messages. The attack affects the asynchronous behaviour of multiple generator rotors, and thus re-closing the affected transmission lines at the proper time is required. However, the main limitation of this work is that the mitigation system is evaluated in a non-realistic scenario, where the attack is simulated through dynamic equations of the power system under cyber-attack.

Jokar and Leun [56] present a specification-based IPS for smart grid applications which use ZigBee-based home area networks. They monitor the network behaviour of different sensor nodes and use Q-learning to implement the prevention response. They extract different network features to learn the pattern of normal communication using Deep Packet Inspection techniques, and evaluate the system carrying out different cyber-attacks, namely (a) radio jamming attacks, (b) stenography attacks, (c) replay attacks, (d) back-off manipulation attacks, (e) DoS against data transmission and (f) DoS against Guaranteed Time Slot requests. However, the IPS is focused on IEEE 802.15.4, and the feature extraction mechanisms for the anomaly detection phase require Deep Packet Inspection, a technique that is not efficient in large networks.

Parras et al. [57] recently proposed the use of Generative Adversarial Imitation Learning, a branch of Inverse RL to mitigate cyberattacks in wireless networks based on CSMA/CA multiple access method. Inverse RL, also known as inverse optimal control, is a learning setting where the learning goal is focused on learning the reward that better explains a given policy of an agent. Therefore, Inverse RL is useful to model an unknown reward function. Parras et al. use Generative Adversarial Imitation Learning, where Deep Neural Networks are used to approximate the agent's reward and the policy. They achieve high detection results and are able to mitigate the back-off attack in both offline and online settings, improving the baseline model results. They argue that their defence mechanism is able to adapt to different RL-based attacks, without assuming any concrete attack type. However, they test the defence mechanisms on backoff attacks only and do not provide results against other types of attacks such as DoS. Finally, another weakness of this general defence mechanism is its computational complexity, especially in online settings which require training the deep neural network multiple times.

Liu et al. propose in [58] an anti-jamming algorithm implemented with deep Q learning with a recursive convolutional neural network to learn the recursive characteristics of the spectrum waterfall information. Spectrum waterfall information is therefore used to obtain observations from the environment. The spectrum waterfall contains both frequency and time data of the network environment. The RL agent learns to select a discretized transmission frequency from a predefined set, with a reward function defined as an SINR-based transmission rate and cost for frequency switching.

Zhang et al. propose in [41] a simplified DNS amplification attack model, mitigating it with a model-free RL agent which observes the traffic load received from a DNS server and the total link load. Even though the paper is not focused on a smart grid domain, the mitigation actions and the model could be easily transferred. The agent implements two different actions: transmitting or discarding all the traffic received in a specified time window. Since the goal of the agent is to eliminate the network congestion after a successful amplification attack, the reward function is designed to avoid the total load of the link goes above a certain threshold: when the load of the link reaches the threshold, a negative reward value is given to the agent. Otherwise, the reward value will be the proportion of legitimate traffic transmitted over the link. Finally, they use a Q-learning method to explore and update the whole state-action space. However, the agent only learns to either drop or allow all the traffic in a specific time window, which can lead to dropping legitimate traffic.

At last, Chen et al. present in [50] a Q-learning-based mitigation strategy for FDI attacks in an automatic voltage controller, replacing the data detected as suspicious with

their own estimation based on maximum likelihood estimation values, in order to enhance the security of the state estimation and the optimal power flow controllers. They model the attacker-defence system as a POMDP, where the attacker learns the optimal attack strategy, and the defender uses a Q-learning algorithm with nearest sequence memory.

Table 2. Strengths and limitations of the discussed works on using RL to defend smart grids.

RL Algorithm	Attack Type	Learning Goal	Strengths (+) and Limitations (−)
DQDN [52]	False Data Injection	Optimize defence strategy by quantifying the observation space with a sliding window	(+) Improves previous approaches in several IEEE bus systems (−) Only considers the physical layer (−) Time complexity grows exponentially with the number of devices (−) Non-realistic scenario, lack of advanced simulation or complex Smart Grid
DPDG [53]	DDoS	Monitor the network load to drop excess traffic to maximize the available bandwidth in the network	(+) Efficient monitoring scheme (−) Not focused on Smart Grid
DPDG [55]	False Data Injection	Find the optimal re-close transmission time after an attack	(+) The implemented attacks affect the asynchronous behaviour of the generator rotors (−) Only acts when the attack already happened (−) Simulated attack scenario with dynamic equations of the power system (−) Only considers the physical layer
Actor-critic NN [51]	Multiple	Learn the optimal strategy to timely defend a CPS by observing the state of the CPS at the cyber and physical level	(+) Real-time operation (+) Learns optimal defence and worst attack policies (−) Non-realistic scenario, lack of advanced simulation or complex Smart Grid (−) Only considers the physical layer
POMDP [54]	DoS FDI/Jamming topology attacks	Lessen its detection delay and false alarm rate by choosing the optimal actions	(+) Model-free algorithm that is able to work online in real time (−) Only considers two actions, continue operation or stop the system (−) Fails to mitigate the attack without stopping the system (−) Their approach does not distinguish between cyberattacks and other kinds of anomalies
Q-learning [56]	Multiple	Learn the optimal actions to mitigate different cyberattacks in ZigBee home area networks	(+) Six different attacks evaluated (+) Combines detection and prevention using different ML-based techniques (−) Focused on IEEE 802.15.4 (−) Requires deep packet inspection
Inverse RL [57]	Back-off attack	Novel general defence mechanism based on Generative Adversarial Imitation Learning	(+) Good performance against back-off attacks (+) Evaluated in both offline and online settings (+) Combines detection and prevention (−) Evaluated only against back-off attack (−) Potentially high computational complexity on online settings
DQN [58]	Jamming	Learn the recursive characteristics of the spectrum waterfall, optimizing the discretized transmission frequency	(+) Considers SINR-based transmission rate and the cost for frequency switching (+) Reduced average detection time and false alarm rate (−) Not focused on Smart Grid
Q-learning [41]	DoS amplification	Learn the optimal actions to eliminate network congestion in a DNS server after an amplification attack	(+) Realistic DNS amplification attack (−) The agent only learns to either transmit or drop all the traffic in a specified time window (−) Not focused on Smart Grid
POMDP [50]	FDI	Maintain optimal power flow	(+) Enhancing the grid resilience with MLE (−) Only considers the physical layer

4. Conclusions and Prospects for Future Research

This paper presented a review of the state of the art on DoS attacks, detection and mitigation techniques in the smart grid. Due to the introduction of information and communication technologies in the grid, the attack surface has been increased, making the grid vulnerable to a wide range of cyberattacks. In particular, DoS attacks targeting different protocols, applications or layers of the smart grid might have serious consequences both on the communication network and on the physical level.

We performed a review of the most important DoS attacks against the smart grid. We studied both the attack techniques at a general level and their particular consequences in the smart grid, identifying the main detection and mitigation actions that might be implemented to prevent them.

Focusing on the detection and mitigation techniques, we performed a review of the most recent works on RL techniques to mitigate DoS attacks in the smart grid. Most of the works on DoS cyberattack detection and mitigation with RL have focused on solving different optimization problems in the smart grid at the control plane. However, there is limited work which tackles the use of RL to mitigate cyberattacks in the smart grid; in particular, there is a lack of holistic approaches which consider and try to mitigate DoS cyberattacks at the communication network and the physical layer at the same time. In particular, the effect of DoS attacks in the smart grid, and how to mitigate these kinds of attacks in near real-time in a realistic simulation is barely investigated.

The main reason behind the limited work in deep RL for smart grid cybersecurity is that deep RL algorithms take a long time to converge, and require a simulation environment

for training and validation. For the deep RL to be applicable to real scenarios, the simulation environment has to be realistic enough so the trained algorithm can be applied to a real environment with confidence [59]. Therefore, most scientific work is focused on small-scale simulations and specific attacks, where the effect of the cyberattack is introduced in the simulation and its effects are investigated at the control plane. To be able to study the effect of DoS attacks in the smart grid at both the communication and physical levels, advanced co-simulation environments which simulate both environments at the same time are required.

There is a clear need for holistic approaches which consider the detection and mitigation of cyberattacks both at the network and control plane in order to effectively mitigate the attack in a timely and accurate manner. In particular, Intrusion Prevention Systems which take into consideration the specific needs of the smart grid need to be studied, designed and evaluated in a realistic smart grid co-simulated environment. The co-simulation environment should have the ability to launch advanced cyberattacks that impact both the communication network and the physical layer, in order to study the best defence mechanisms that can mitigate advanced attacks at both levels.

Author Contributions: Conceptualization, I.O.-F. and F.L.; software, not applicable; validation, I.O.-F. and F.L.; formal analysis, I.O.-F. and F.L.; investigation, I.O.-F.; resources, I.O.-F.; data curation, not applicable; writing—original draft preparation, I.O.-F.; writing—review and editing, I.O.-F. and F.L.; visualization, I.O.-F.; supervision, I.O.-F. and F.L.; project administration, I.O.-F. and F.L.; funding acquisition, I.O.-F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ayudas Cervera para Centros Tecnológicos grant of the Centre for Industrial Technological Development (CDTI) under the project ÉGIDA (CER-20191012).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CPS	Cyber–Physical System
DDPG	Deep Deterministic Policy Gradient
DDoS	Distributed Denial of Service
DNS	Domain Name System
DoS	Denial of Service
DQN	Deep Q Learning
FDI	False Data Injection
GPS	Global Positioning System
IPS	Intrusion Prevention System
NIDS	Network Intrusion Detection Systems
NTP	Network Time Protocol
POMDP	Partially observable Markov decision process
PTP	Precision Time Protocol
RL	Reinforcement Learning
SCADA	Supervisory Control and Data Acquisition
SDN	Software Defined Networks
SNMP	Simple Network Management Protocol

References

1. Pham, L.N.H. Exploring Cyber-Physical Energy and Power System: Concepts, Applications, Challenges, and Simulation Approaches. *Energies* **2023**, *16*, 42. [CrossRef]
2. Fang, X.; Misra, S.; Xue, G.; Yang, D. Smart Grid—The New and Improved Power Grid: A Survey. *IEEE Commun. Surv. Tutor.* **2012**, *14*, 944–980. [CrossRef]

3. Radoglou-Grammatikis, P.I.; Sarigiannidis, P.G. Securing the Smart Grid: A Comprehensive Compilation of Intrusion Detection and Prevention Systems. *IEEE Access* **2019**, *7*, 46595–46620. [CrossRef]
4. Pillitteri, V.Y.; Brewer, T.L. *Guidelines for Smart Grid Cybersecurity*; NIST: Gaithersburg, MD, USA, 2014.
5. Huseinović, A.; Mrdović, S.; Bicakci, K.; Uludag, S. A survey of denial-of-service attacks and solutions in the smart grid. *IEEE Access* **2020**, *8*, 177447–177470. [CrossRef]
6. Cheng, Z.; Yue, D.; Hu, S.; Huang, C.; Dou, C.; Chen, L. Resilient load frequency control design: DoS attacks against additional control loop. *Int. J. Electr. Power Energy Syst.* **2020**, *115*, 105496. [CrossRef]
7. Fekete, B.M.; Revenga, C.; Todd, M. The Global Risks Report 2018 13th Edition. Available online: http://www3.weforum.org/docs/WEF_GRR18_Report.pdf (accessed on 15 October 2022).
8. Liu, J.; Xiao, Y.; Li, S.; Liang, W.; Chen, C.P. Cyber security and privacy issues in smart grids. *IEEE Commun. Surv. Tutor.* **2012**, *14*, 981–997. [CrossRef]
9. Goudarzi, A.; Ghayoor, F.; Waseem, M.; Fahad, S.; Traore, I. A Survey on IoT-Enabled Smart Grids: Emerging, Applications, Challenges, and Outlook. *Energies* **2022**, *15*, 6984. [CrossRef]
10. Fares, A.A.Y.R.; de Caldas Filho, F.L.; Giozza, W.F.; Canedo, E.D.; Lopes de Mendonça, F.L.; Amvame Nze, G.D. DoS Attack Prevention on IPS SDN Networks. In Proceedings of the 2019 Workshop on Communication Networks and Power Systems (WCNPS), Brasilia, Brazil, 3–4 October 2019; pp. 1–7. [CrossRef]
11. Raja, D.J.S.; Sriranjani, R.; Parvathy, A.; Hemavathi, N. A Review on Distributed Denial of Service Attack in Smart Grid. In Proceedings of the IEEE 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2–24 June 2022; pp. 812–819.
12. Berthier, R.; Sanders, W.H.; Khurana, H. Intrusion Detection for Advanced Metering Infrastructures: Requirements and Architectural Directions. In Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications, Gaithersburg, MD, USA, 4–6 October 2010; pp. 350–355. [CrossRef]
13. Cisco, I. *NetFlow Configuration Guide Release 12.4*; Cisco Documentation; Cisco Systems: San Jose, CA, USA, 2007.
14. Quittek, J.; Zseby, T.; Claise, B.; Zander, S. Requirements for IP Flow Information Export (IPFIX); RFC Editor, October 2004. Available online: https://www.rfc-editor.org/info/rfc3917 (accessed on 15 October 2022).
15. Sperotto, A.; Schaffrath, G.; Sadre, R.; Morariu, C.; Pras, A.; Stiller, B. An overview of IP flow-based intrusion detection. *IEEE Commun. Surv. Tutor.* **2010**, *12*, 343–356. [CrossRef]
16. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [CrossRef]
17. Asri, S.; Pranggono, B. Impact of distributed denial-of-service attack on advanced metering infrastructure. *Wirel. Pers. Commun.* **2015**, *83*, 2211–2223. [CrossRef]
18. Groat, S.; Dunlop, M.; Urbanksi, W.; Marchany, R.; Tront, J. Using an IPv6 moving target defense to protect the Smart Grid. In Proceedings of the 2012 IEEE PES Innovative Smart Grid Technologies (ISGT), Washington, DC, USA, 16–20 January 2012; pp. 1–7. [CrossRef]
19. Choi, K.; Chen, X.; Li, S.; Kim, M.; Chae, K.; Na, J. Intrusion Detection of NSM Based DoS Attacks Using Data Mining in Smart Grid. *Energies* **2012**, *5*, 4091–4109. [CrossRef]
20. Jin, D.; Nicol, D.M.; Yan, G. An event buffer flooding attack in DNP3 controlled SCADA systems. In Proceedings of the 2011 Winter Simulation Conference (WSC), Phoenix, AZ, USA, 11–14 December 2011; pp. 2614–2626. [CrossRef]
21. Zhang, F.; Mahler, M.; Li, Q. Flooding attacks against secure time-critical communications in the power grid. In Proceedings of the 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), Dresden, Germany, 23–27 October 2017; pp. 449–454. [CrossRef]
22. Li, Q.; Ross, C.; Yang, J.; Di, J.; Balda, J.C.; Mantooth, H.A. The effects of flooding attacks on time-critical communications in the smart grid. In Proceedings of the 2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 18–20 February 2015; pp. 1–5. [CrossRef]
23. Pelechrinis, K.; Iliofotou, M.; Krishnamurthy, S.V. Denial of Service Attacks in Wireless Networks: The Case of Jammers. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 245–257. [CrossRef]
24. Temple, W.G.; Chen, B.; Tippenhauer, N.O. Delay makes a difference: Smart grid resilience under remote meter disconnect attack. In Proceedings of the 2013 IEEE International Conference on Smart Grid Communications (SmartGridComm), Vancouver, BC, Canada, 21–24 October 2013; pp. 462–467. [CrossRef]
25. Li, H.; Lai, L.; Qiu, R.C. A denial-of-service jamming game for remote state monitoring in smart grid. In Proceedings of the 2011 45th Annual Conference on Information Sciences and Systems, Baltimore, MD, USA, 23–25 March 2011; pp. 1–6. [CrossRef]
26. Chatfield, B.; Haddad, R.J.; Chen, L. Low-Computational Complexity Intrusion Detection System for Jamming Attacks in Smart Grids. In Proceedings of the 2018 International Conference on Computing, Networking and Communications (ICNC), Maui, HI, USA, 5–8 March 2018; pp. 367–371. [CrossRef]
27. Zhang, Z.; Gong, S.; Dimitrovski, A.D.; Li, H. Time Synchronization Attack in Smart Grid: Impact and Analysis. *IEEE Trans. Smart Grid* **2013**, *4*, 87–98. [CrossRef]
28. Gaderer, G.; Treytl, A.; Sauter, T. Security aspects for IEEE 1588 based clock synchronization protocols. In Proceedings of the 2006 IEEE International Workshop on Factory Communication Systems, Turin, Italy, 28–30 June 2006; pp. 247–250. [CrossRef]
29. Moussa, B.; Debbabi, M.; Assi, C. A Detection and Mitigation Model for PTP Delay Attack in an IEC 61850 Substation. *IEEE Trans. Smart Grid* **2018**, *9*, 3954–3965. [CrossRef]

30. Yang, Y.; McLaughlin, K.; Sezer, S.; Littler, T.; Pranggono, B.; Brogan, P.; Wang, H. Intrusion detection system for network security in synchrophasor systems. In Proceedings of the IET International Conference on Information and Communications Technologies, Beijing, China, 27–29 April 2013.
31. Vuković, O.; Dán, G. Security of Fully Distributed Power System State Estimation: Detection and Mitigation of Data Integrity Attacks. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1500–1508. [CrossRef]
32. Chen, P.Y.; Yang, S.; McCann, J.A.; Lin, J.; Yang, X. Detection of false data injection attacks in smart-grid systems. *IEEE Commun. Mag.* **2015**, *53*, 206–213. [CrossRef]
33. Zhang, Y.; Wang, J.; Chen, B. Detecting False Data Injection Attacks in Smart Grids: A Semi-Supervised Deep Learning Approach. *IEEE Trans. Smart Grid* **2021**, *12*, 623–634. [CrossRef]
34. Bogdanoski, M.; Suminoski, T.; Risteski, A. Analysis of the SYN flood DoS attack. *Int. J. Comput. Netw. Inf. Secur.* **2013**, *5*, 1–11. [CrossRef]
35. Gupta, N.; Jain, A.; Saini, P.; Gupta, V. DDoS attack algorithm using ICMP flood. In Proceedings of the IEEE 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 4082–4084.
36. Huseinovic, A.; Mrdovic, S.; Bicakci, K.; Uludag, S. A Taxonomy of the Emerging Denial-of-Service Attacks in the Smart Grid and Countermeasures. In Proceedings of the 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 20–21 November 2018; pp. 1–4. [CrossRef]
37. Eidson, J.C.; Fischer, M.; White, J. IEEE-1588 Standard for a precision clock synchronization protocol for networked measurement and control systems. In Proceedings of the 34th Annual Precise Time and Time Interval Systems and Applications Meeting, Reston, VA, USA, 3–5 December 2002; pp. 243–254.
38. Fan, Y.; Zhang, Z.; Trinkle, M.; Dimitrovski, A.D.; Song, J.B.; Li, H. A Cross-Layer Defense Mechanism Against GPS Spoofing Attacks on PMUs in Smart Grids. *IEEE Trans. Smart Grid* **2015**, *6*, 2659–2668. [CrossRef]
39. Baumgartner, B.; Riesch, C.; Schenk, W. The impact of gps vulnerabilities on the electric power grid. In Proceedings of the XX IMEKO TC–4 International Symposium on Research on Electrical and Electronic Measurement for the Economic Upturn, Benevento, Italy, 15–17 September 2014; pp. 183–188.
40. Anagnostopoulos, M. Amplification DoS Attacks. In *Encyclopedia of Cryptography, Security and Privacy*; Jajodia, S., Samarati, P., Yung, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 1–3. [CrossRef]
41. Zhang, Y.; Cheng, Y. An Amplification DDoS Attack Defence Mechanism using Reinforcement Learning. In Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 19–23 August 2019; pp. 634–639. [CrossRef]
42. Anagnostopoulos, M.; Kambourakis, G.; Kopanos, P.; Louloudakis, G.; Gritzalis, S. DNS amplification attack revisited. *Comput. Secur.* **2013**, *39*, 475–485. [CrossRef]
43. Gondim, J.J.; de Oliveira Albuquerque, R.; Orozco, A.L.S. Mirror saturation in amplified reflection Distributed Denial of Service: A case of study using SNMP, SSDP, NTP and DNS protocols. *Future Gener. Comput. Syst.* **2020**, *108*, 68–81. [CrossRef]
44. Liang, G.; Weller, S.R.; Zhao, J.; Luo, F.; Dong, Z.Y. The 2015 Ukraine Blackout: Implications for False Data Injection Attacks. *IEEE Trans. Power Syst.* **2017**, *32*, 3317–3318. [CrossRef]
45. Rahman, M.A.; Mohsenian-Rad, H. False data injection attacks with incomplete information against smart power grids. In Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM), Anaheim, CA, USA, 3–7 December 2012; pp. 3153–3158. [CrossRef]
46. Wan, Z.; Li, H.; He, H.; Prokhorov, D. Model-free real-time EV charging scheduling based on deep reinforcement learning. *IEEE Trans. Smart Grid* **2018**, *10*, 5246–5257. [CrossRef]
47. Wan, Z.; Jiang, C.; Fahad, M.; Ni, Z.; Guo, Y.; He, H. Robot-assisted pedestrian regulation based on deep reinforcement learning. *IEEE Trans. Cybern.* **2018**, *50*, 1669–1682. [CrossRef]
48. Duan, J.; Yi, Z.; Shi, D.; Lin, C.; Lu, X.; Wang, Z. Reinforcement-learning-based optimal control of hybrid energy storage systems in hybrid AC–DC microgrids. *IEEE Trans. Ind. Inform.* **2019**, *15*, 5355–5364. [CrossRef]
49. Yan, J.; He, H.; Zhong, X.; Tang, Y. Q-learning-based vulnerability analysis of smart grid against sequential topology attacks. *IEEE Trans. Inf. Forensics Secur.* **2016**, *12*, 200–210. [CrossRef]
50. Chen, Y.; Huang, S.; Liu, F.; Wang, Z.; Sun, X. Evaluation of Reinforcement Learning-Based False Data Injection Attack to Automatic Voltage Control. *IEEE Trans. Smart Grid* **2019**, *10*, 2158–2169. [CrossRef]
51. Feng, M.; Xu, H. Deep reinforecement learning based optimal defense for cyber-physical system in presence of unknown cyber-attack. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017; pp. 1–8. [CrossRef]
52. An, D.; Yang, Q.; Liu, W.; Zhang, Y. Defending against Data Integrity Attacks in Smart Grid: A Deep Reinforcement Learning-Based Approach. *IEEE Access* **2019**, *7*, 110835–110845. [CrossRef]
53. Liu, Y.; Dong, M.; Ota, K.; Li, J.; Wu, J. Deep Reinforcement Learning based Smart Mitigation of DDoS Flooding in Software-Defined Networks. In Proceedings of the 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Barcelona, Spain, 17–19 September 2018; pp. 1–6. [CrossRef]
54. Kurt, M.N.; Ogundijo, O.; Li, C.; Wang, X. Online Cyber-Attack Detection in Smart Grid: A Reinforcement Learning Approach. *IEEE Trans. Smart Grid* **2019**, *10*, 5174–5185. [CrossRef]

55. Wei, F.; Wan, Z.; He, H. Cyber-Attack Recovery Strategy for Smart Grid Based on Deep Reinforcement Learning. *IEEE Trans. Smart Grid* **2020**, *11*, 2476–2486. [CrossRef]
56. Jokar, P.; Leung, V.C.M. Intrusion Detection and Prevention for ZigBee-Based Home Area Networks in Smart Grids. *IEEE Trans. Smart Grid* **2018**, *9*, 1800–1811. [CrossRef]
57. Parras, J.; Almodóvar, A.; Apellániz, P.A.; Zazo, S. Inverse Reinforcement Learning: A New Framework to Mitigate an Intelligent Backoff Attack. *IEEE Internet Things J.* **2022**, *9*, 24790–24799. [CrossRef]
58. Liu, X.; Xu, Y.; Jia, L.; Wu, Q.; Anpalagan, A. Anti-jamming communications using spectrum waterfall: A deep reinforcement learning approach. *IEEE Commun. Lett.* **2018**, *22*, 998–1001. [CrossRef]
59. Zhang, D.; Han, X.; Deng, C. Review on the research and practice of deep learning and reinforcement learning in smart grids. *CSEE J. Power Energy Syst.* **2018**, *4*, 362–370. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.