Ad Hoc Networks 164 (2024) 103642

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Survey paper

A survey on security of UAV and deep reinforcement learning

Burcu Sönmez Sarıkaya, Şerif Bahtiyar [*]

Cyber Security and Privacy Research Lab, Department of Computer Engineering, Istanbul Technical University, Istanbul, Maslak, 34469, Türkiye

ARTICLE INFO

Keywords:
Unmanned aerial vehicles
Security
Machine learning
Deep reinforcement learning
Threat
Attack
Vulnerability
Countermeasure

ABSTRACT

Recently, the use of unmanned aerial vehicles (UAV)s for accomplishing various tasks has gained a significant interest from both civilian and military organizations due to their adaptive, autonomous, and flexibility nature in different environments. The characteristics of UAV systems introduce new threats from which cyber attacks may benefit. Adaptive security solutions for UAVs are required to counter the growing threat surface. The security of UAV systems has therefore become one of the fastest growing research topics. Machine learning based security mechanisms have a potential to provide effective countermeasures that complement traditional security mechanisms. The main motivation of this survey is to the lack of a comprehensive literature review about reinforcement learning based security solutions for UAV systems. In this paper, we present a comprehensive review on the security of UAV systems focusing on deep-reinforcement learning-based security solutions. We present a general architecture of an UAV system that includes communication systems to show potential sources of vulnerabilities. Then, the threat surface of UAV systems is explored. We explain attacks on UAV systems according to the threats in a systematic way. In addition, we present countermeasures in the literature for each attack on UAVs. Furthermore, traditional defense mechanisms are explained to highlight requirements for reinforcement based security solutions on UAVs. Next, we present the main reinforcement algorithms. We examine security solutions with reinforcement learning algorithms and their limitations in a holistic approach. We also identify research challenges about reinforcement based security solutions on UAVs. Briefly, this survey provides key guidelines on UAV systems, threats, attacks, reinforcement learning algorithms, the security of UAV systems, and research challenges.

1. Introduction

Unmanned Aerial Vehicles (UAVs) are aircraft that are remotely operated by human operators or autonomously controlled by onboard computers. They have become increasingly popular in various civilian and military applications due to their versatility, cost-effectiveness, and ability to operate in hazardous or inaccessible environments. As with any connected device, UAVs are vulnerable to cyber attacks that can compromise their functionality, data security, and safety. UAVs use a combination of communication networks, sensors, and software systems. UAVs collect and process data, control their flight, and carry out specific missions. These systems generate and exchange large amounts of sensitive information, such as location data, video footage, sensor data, and control commands. Attackers can exploit vulnerabilities in these systems to gain unauthorized access, manipulate data, disrupt communications, hijack systems' control, or cause physical damages. Thus, attacks may compromise security goals of UAVs, which are known as confidentiality, integrity, and availability [1–3].

Communication requirements of UAVs depend on the environment which determines the security requirements of UAVs. The complex communication environment of UAVs extends the attack surface. Therefore, it is necessary to apply multiple security mechanisms at the same time to ensure the security of UAVs. This circumstance necessitates the design of new security systems. The emerging security requirements also increase the communication cost of UAVs. Thus, the main challenge is to create a new attack detection system that supports dynamic environmental conditions for UAVs.

UAVs interact with many networks, therefore, security solutions of the networks should be addressed to cope with the grand challenge. Emerging technologies such as blockchain, software-defined networks (SDN), machine learning, fog, and edge computing have been explored in [4,5] as part of the solution architectures. Considering the resource constraints of drones, some blockchain-based systems [6–8] may provide a solution instead of using all properties of heavy resource required chains. SDN technology may ensure network reliability by allowing the controller to closely monitor data traffic in UAVs [9]. Fog computing may be used to maintain computing capabilities near drones without exceeding their capacity [10].

[*] Corresponding author.
E-mail addresses: sonmezb18@itu.edu.tr (B.S. Sarıkaya), bahtiyars@itu.edu.tr (Ş. Bahtiyar).

https://doi.org/10.1016/j.adhoc.2024.103642
Received 23 June 2024; Received in revised form 20 August 2024; Accepted 27 August 2024
Available online 30 August 2024

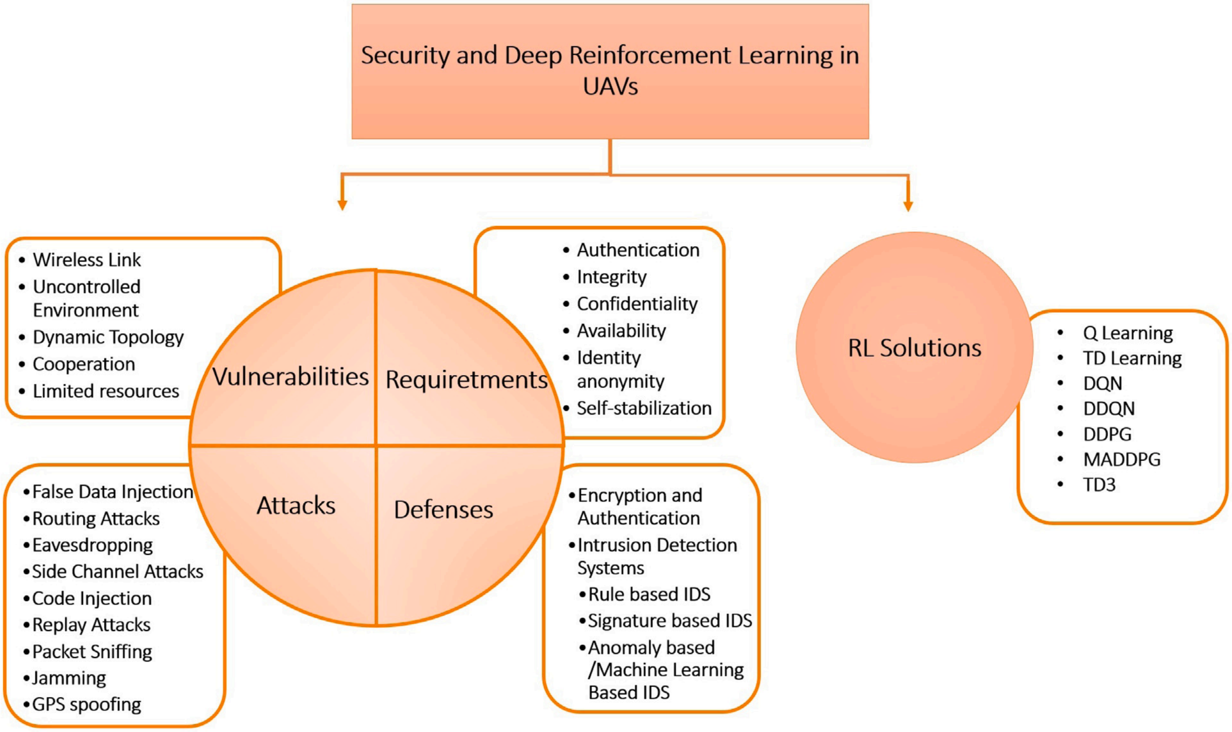


Fig. 1. Security and reinforcement learning in UAVs.

Reinforcement learning (RL) is a subfield of machine learning that brings a reward-driven behavior from psychology to artificial intelligence that creates a powerful alternative to regular controllers [11]. RL stands out from usual control methods since it only requires a goal definition to work and no prior knowledge about the environment or system is required. Actions to reach the goal are chosen by RL algorithms based on positive or negative outcomes of previous interactions with the environment.

RL maps an appropriate action to each state. After performing the assigned action, a positive or negative feedback increases or decreases the chance of the action being chosen again. Updating mappings iteratively, RL finds best actions for each state to reach the goal as soon as possible. Exploitation of experiences to select better actions among candidates eliminates the requirement of the system or environment models that are given in advance [12].

Deep reinforcement learning (DRL) is a promising technique that will enhance the security of UAVs and communication networks. DRL helps UAVs to learn from experiences and make decisions based on the environmental feedback. On the other hand, DRL poses several technical and ethical challenges that need to be addressed. Interpretation of learned policies, data privacy, and the potential for adversarial attacks are possible [13]. For example, several RL applications require data accesses, sensing the environment and collecting data, resource allocations for wireless connectivity, UAV-enabled mobile edge computing data, localization data, trajectory planning, and network security in Multi-UAV wireless networks. Nevertheless, the absence of specific models for security that support RL-based solutions for UAV security requires further research. Our motivation in this research is the lack of a comprehensive review in the literature on deep reinforcement learning applications and UAV security.

In this research, we propose a novel survey on deep reinforcement learning solutions to secure UAVs since traditional security methods for such vehicles require a large amount of computational resources and are impractical to implement in a real-time manner [14]. Our contributions in this research are as follows.

- We present a comprehensive review of reinforcement learning techniques that provide security for UAVs.
- We provide a taxonomy of existing solutions for the most relevant tasks studied as a part of security mechanisms for UAVs, such as vulnerabilities, security requirements, cyber-attacks, and defenses.
- We analyze potential solutions to reduce the effects of attacks by using security controls and rules with deep reinforcement learning for early detections of attacks.
- We emphasize practical DRL solutions for security of UAVs.
- We elaborate on open problems and future research directions about the security of UAVs with DRL.

The rest of the paper is organized as follows. Section 2 is about UAVs and their security requirements. The next section includes threats and attacks in UAVs. Section 4 contains traditional defense mechanisms applied on UAVs. Section 5 includes a detailed investigation of recent research activities on RL-based solutions in UAV security. Then, analyses of security approaches are presented in the next section. We discuss limitations of RL-based solutions and research challenges in Section 7. The last section is devoted to the conclusion. The general structure of the paper is shown in Fig. 1.

2. UAV and security

2.1. UAV systems and communications

The development of UAV technology has led to the creation of different types of drones with different shapes and weights. This section examines the general system of UAVs and their communication infrastructures that are vulnerable to cyber-attack. The investigation focuses on the high-level structure and fundamental components of a UAV system. These include the communications infrastructure of UAVs, the Ground Control Station (GCS), and satellite communications technologies.

In general, an UAV system consists of an unmanned aircraft, a ground control station and a communications data link [15,16]. The

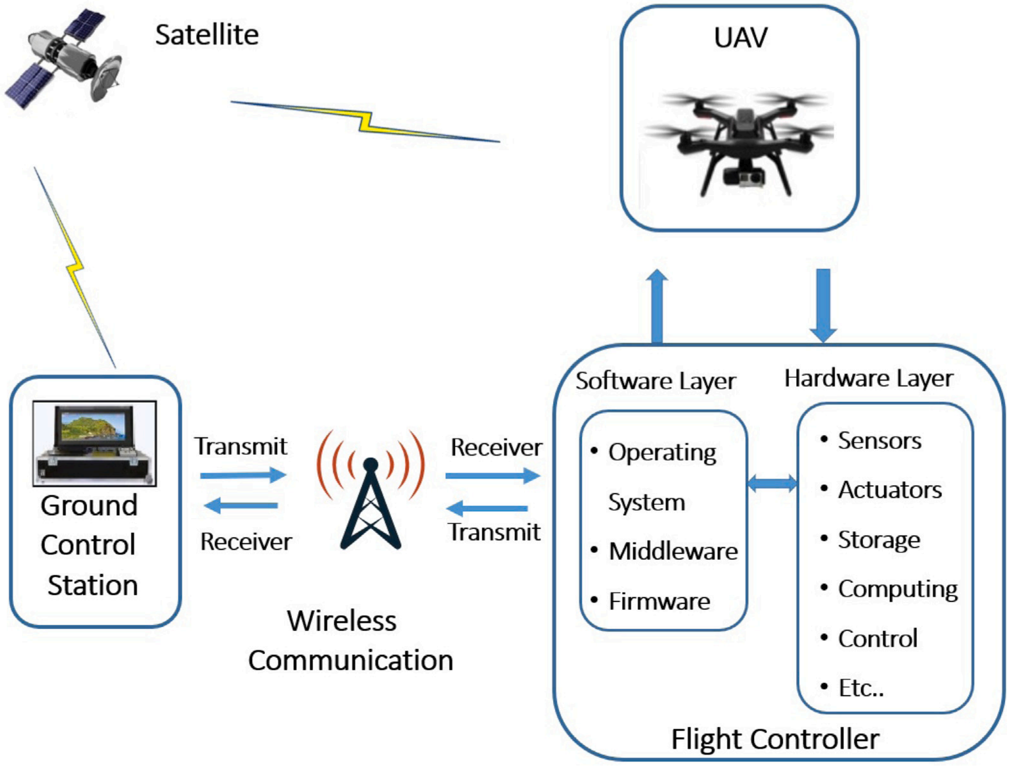


Fig. 2. General architecture of an UAV system.

unmanned aircraft represents the core of an unmanned aerial system and consists of an airframe, a propulsion system, a flight controller, a precision navigation system, and a sense-and-avoid system [17]. In this research, we consider security of UAVs so we consider only the corresponding building blocks that will create a vulnerability in the system. For instance, an aircraft's flight controller, communication lines, and sensors are blocks of UAVs that may contain vulnerabilities. The high-level architecture of a UAV system and its main elements are shown in Fig. 2.

The flight controller is the central processing unit of a drone that is responsible for stabilizing the aircraft during flight and collecting data from sensors. At the hardware level, the flight controller configuration includes rechargeable batteries, actuators, and various sensors, such as GPS and accelerators, as well as a wireless communication module. The software architecture consists of three layers, namely, firmware, middleware, and operating system. The firmware issues machine code instructions, the middleware manages communication between services and the operating system, which is often an RTOS that handles real-time data processing [18].

The flight controller makes communications with the ground control station easy. The controller processes commands and translates them into actions for actuators. Telemetric signals are transmitted to GCS via multiple channels. The flight controller may integrate sensors or may communicate with external units to enhance data acquisition capabilities, making it a critical component of the UAV system's distributed embedded architecture [17] .

UAVs are not a fully self-acting artificial intelligence products, but they are a part of the whole of the communication infrastructure. For this reason, an UAV has to communicate with other UAVs and ground control stations to complete its mission. The communication lines in UAVs are shown in Fig. 3. Two types of methods are used for the communication of UAVs [19]. The first method is to check the movement of an UAV with signal communication. The other is the transfer of data related to the task that the UAV should complete with data communication.

An UAV collects, processes, and transmits sensitive data according to its usage area. Therefore, the security of data that may be related to strategic operations, environmental surveillance, and communication is very important. The expanding application domains of UAVs, the huge communication options, and many implementations with different infrastructures, such as wireless sensor networks, mobile ad hoc networks (MANETs), Ad-Hoc Network (VANET), Flying Ad-Hoc Network (FANET), and GPS make UAVs open to security attacks [20–23]. In addition, UAVs are used to communicate with each other with less energy-consuming EMA-style protocols. Bluetooth, Zigbee, and WiFi protocols are used for short distances between UAVs and GCS. WiMAX and Cellular protocols are used for long distances. GPS coordinates are used for the communication of UAVs with satellites, usually WiMAX and Cellular communications [24].

UAVs have different network technologies for communicating with multiple UAVs (UAV-to-UAV, U2U) and between UAVs and other systems (UAV-to-Infrastructure: U2I), such as GCS, WSN, and ground reactors as shown in Fig. 3. Ad-hoc network is recommended as a good solution. [25] contains key issues related to UAV communication networks that include characteristics of existing ad hoc networks to classify UAV networks by topology. [26] focuses on communication between an UAV and a GCS. They are classified technologies for communication links, such as Zigbee wimax, IEEE 802.xx, by the range and the data rate. The most important problem in ensuring cooperation between Multi-UAVs is the communication infrastructure of UAVs that requires network supports and service components. A categorization of security risks and solutions related to FANETs and UAV communications is presented according to the four OSI layers in [23].

The most significant attack vector of UAVs is its communication. Specifically, it is critical for UAVs to secure wireless channels that are vulnerable to attacks. Obviously, UAVs communicate with each other and through the ground control station via wireless channels that are open to various attacks. In fact, launching an attack on UAVs is very easy. Special and critical UAV information can be accessed by unauthorized users. On the other hand, due to the lack of security

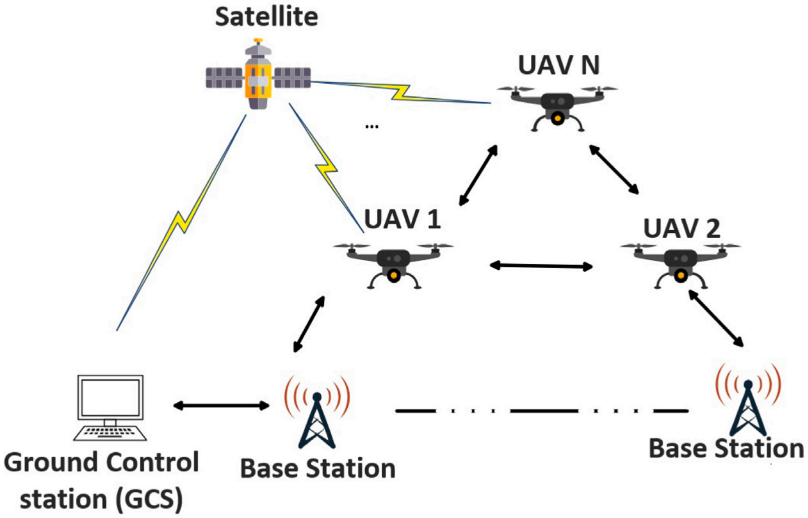


Fig. 3. The communication links of UAVs.

mechanisms in the applications of satellite-connected UAVs, legal accesses to the essential services may be prevented, which is a kind of denial of service (DoS) attack. Thus, UAVs require security mechanism to preserve confidentiality, integrity, and availability of whole system in an UAV. Communications of UAVs are done by using data transfer or more specifically by using signal communications. The main security requirements for data transfer in a communication are integrity and confidentiality. The security requirement for signal processing is accessibility, which is sometimes known as availability.

Attacks on UAVs may endanger confidentiality, integrity, and availability of data [27] that are security requirements of the UAV system cyber security threat model [28]. There are also cyber–physical attacks on UAVs that have targets on the UAV network layers [29]. Attackers targeting confidentiality have unauthorized access to the system by capturing sensitive data or sensor data. To this end, the attacker can listen to the communication between the UAV and the GCS, then transmit data to other malicious actors. It may threaten security by collecting intelligence about the task of UAVs. Such attacks are referred as eavesdropping attacks in the literature [30].

Encryption methods are used to prevent attacks. For instance, data on wireless communication lines must be encrypted [31]. Attacks may disrupt integrity of data in the communication network that causes unauthorized modification of components and tasks of UAVs. Examples of such attacks targeting UAV data integrity are replay attacks [32]. To prevent these attacks, hash functions, such as a message authentication code, are used that check whether data are corrupted.

UAVs are subject to different attacks according to changing environments, such as the network structure, coverage areas and communication channels. Specifically, an UAV is open to different types of cyber attacks, such as jamming, spoofing, replay attacks, due to the changing communication environments of the UAV and wireless channels. Malicious attacks may interrupt communications to reduce the availability of UAVs. Recently, the jamming and GPS spoofing attacks on the wireless sensor network are used to interrupt communications among UAVs. In order to prevent these attacks, UAVs need a variable and adaptive security mechanism. In particular, it is very important to ensure security in the communicating area, which is critical to complete the task of UAVs. Therefore, jamming and GPS spoofing attacks, which consider the availability of UAVs and cause communication breakdowns, are investigated by many researchers [33].

In order to provide security against attacks on UAVs, an independent method of the environment is necessary to quickly detect attacks. At this point, deep-reinforcement methods provide remarkable options. DRL is suitable and the fastest route for an UAV that provides an independent environment for self-adaptation against attacks. An important summary of deep reinforcement applications in communication networks is given in [34]. Thus, understanding and optimizing UAV communications systems is crucial for achieving secure operations with UAVs in a variety of applications.

2.2. Security requirements of UAV

UAVs are becoming increasingly important in various industries, including military, agriculture, research, etc. UAVs must have strong security because they process sensitive data or operate in mission-critical infrastructures. The basic security requirements for UAVs are authentication, integrity, confidentiality, availability, identity anonymity, and self-stabilization.

Authentication is very important for all nodes in the infrastructure and messages passing through the UAV network. This requirement ensures that only authenticated nodes may participate in the routing process. An attacker can spoof a legal node, gain access to confidential information, and even interfere with network operation without authentication [35–37].

Integrity includes the consistency, accuracy, and reliability of data packets. It ensures that the attacker does not alter data on the traffic or GPS coordinates by adding, deleting, or modifying the data transmission. If the integrity mechanism of UAV network is weak, the integrity of the entire UAVs may be at risk [35–37].

Confidentiality is the measure used to prevent sensitive information from being visible to a wrong node or malicious node. It assures that data is visible only to the right node. Confidentiality ensures that UAV network, payload traffic, command and control traffic, and sensitive information are not visible to unauthorized nodes. In the absence of a privacy mechanism, an attacker can launch attacks by improperly collecting sensitive information about the target [35–38].

Availability ensures that all services provided by an UAV system are always available to authorized entities or devices within an UAV system, even at the time of the attack. It ensures that the network has functional and useful information, such as command and traffic control during UAV network activity. In reality, this is difficult to achieve due to the limitations in delays and the critical nature of an UAV mission [35–37].

Identity anonymity ensures that an attacker cannot obtain the true identity of users from intercepted messages in the communication of an UAV network [36]. This requirement may be implemented as a part of authentication system within an UAV system. Both user authentication

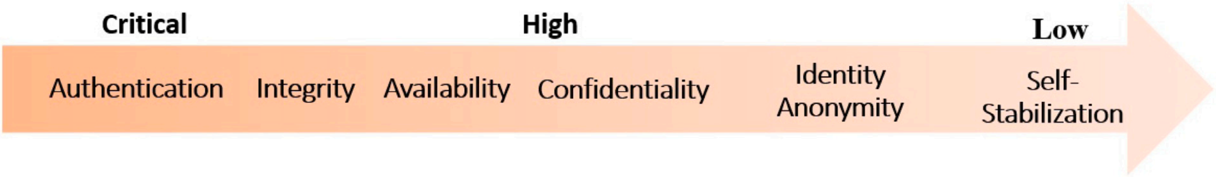


Fig. 4. Security requirements for UAVs.

and packet authentication may be needed to be implemented for UAVs that have critical missions.

Self-stabilization in UAV network communication protocols must be able to automatically recover from any attack without the need of human intervention. This requirement ensures that malicious packets do not permanently damage the network [36]. DRL is a good candidate to provide self-stabilization requirement for UAVs.

The security requirements for UAVs contain specific challenges to be implemented, as the vulnerabilities associated with UAV operations are diverse. It is critical to ensure that only authorized people can access and control UAVs. Authentication mechanisms must be robust to prevent unauthorized access. The integrity of data and instructions transmitted between the ground control station and the UAVs is very critical. Interference with commands or data must be avoided. Availability for UAVs is also critical, which prevents DoS attacks. The prioritization of requirements can vary depending on specific use cases, regulations, and the operational context of the UAV. Regular risk assessments and updates on security measures are essential to counteract evolving threats. In general, prioritized security requirements of UAVs are shown in Fig. 4.

3. Threats and attacks on UAV

3.1. Security threats

While unmanned aerial vehicles offer innovative capabilities, they are prone to vulnerabilities. Understanding these vulnerabilities is critical to ensuring the security and reliability of UAV operations. This section provides an in-depth analysis of various threats and corresponding vulnerabilities in UAVs.

Wireless Link in UAV networks uses wireless connections to send and receive radio signals. In general, anyone can listen a frequency and receive signals with antennas configured for specific frequencies, such as GPS signals sent by GCS and UAVs. Specifically, GCS gives UAVs real-time control over the communications link to a satellite (uplink) and a command traffic. It is sufficient for attackers to simply detect the signal and generate a noise signal in radio communications. [39] contains an overview of applying machine learning techniques to tackle fundamental challenges in wireless communications within the Internet of Things (IoT) with the specific focus on the ad-hoc networking dimension.

A Wi-fi-based communication may have vulnerabilities. If a Wi-fi-based attack is carried out successfully and the control station does not react quickly, the UAV will not be able to receive any command for a certain period of time. This may cause an attacker to hijack an UAV. Due to data connectivity capabilities, wireless connections of UAVs often have lower bandwidth than wired networks. An attacker can exploit this feature by sending fake packets to UAVs. These packets may consume bandwidth and interfere with normal communication [40].

The use of radio transmission combined with a small size, a low cost, and a limited power makes the wireless link more susceptible to denial of service attacks. Wireless connectivity allows attackers to actively or passively eavesdrop on communication. Attacks on a wireless connection can come from all directions and the node can be captured, which may cause leakage of sensitive information [41].

Uncontrolled Environment in UAV networks does not have a central authority such as a switch management system to handle incoming and outgoing network packets, as in wired networks. Due to the lack of a key management system, inside and outside attacks on the network are possible. An attacker can send and receive data traffic while it is within the transmission range of an UAV [35].

Dynamic Topology is another problem that arises with an UAV network, where the environment prevents precisely to detect the enemy node. It is often difficult to distinguish between a malfunctioning node due to the dynamic topology of UAVs and a legal node that appears to be faulty because of the poor link quality. For instance, routing algorithms based on a topology prediction and a self-adaptive learning, which are specifically designed to adapt to the dynamic topology of the network, have been investigated in [42]. Future trends in security and privacy are discussed in addition to routing, connectivity, topology control, and energy efficiency. The difference in routing cost in both cases can be quite small. Therefore, such attacks are difficult to resolve [43].

Cooperation of routing algorithms for an UAV network require all nodes to participate in the discovery of the topology and the transmission of data. Security transactions between nodes are ignored. When a node searches for a route, it broadcasts the message regardless of the recipient's identity. Thus, an attacker could easily participate in the routing process and it can damage the UAV network topology by filtering or blocking the control traffic [35].

Limited Resources in UAVs represent limited information processing and storage capabilities, which are depending on the size of an UAV. By exploiting this vulnerability of the UAV, an attacker can launch an attack that aims to drain the UAV's resources, such as the battery. The implemented security solution directly dependents on the power and the storage capacity of UAVs. For example, the memory of an UAV must be large enough to hold all the variables needed to run asymmetric cryptographic algorithms. For low-power UAVs, this may be difficult because digital signature-based authentication methods require high computation power [44].

In the literature, there are solutions that consider vulnerabilities in different parts of UAV systems, such as routing, partition, connectivity, recovery, fault detection, dynamic mobility, unstable paths, power control, resource allocation, path planning, energy consumption, and communication overhead. For example, a k-means online learning routing protocol (KMORP) is specifically designed for ad hoc UAV networks with a Markov mobility model [45]. KMORP is particularly effective for specialized UAV networks as it swiftly adapts to dynamic changes in the network environment, such as UAV mobility, interference, and signal degradation. This approach guarantees optimal data transmissions and communications by expertly adapting to variations. KMORP utilizes clustering to reduce communication overheads between nodes, which help to alleviate the network's overall load.

An approach called Unmanned Aerial Vehicles (UAV)-assisted Network Segmentation Detection and Connection Restoration (UAV-NetRest) that covers all stages from a network fault detection to a partition resolution is presented in [46]. The speed of a fault detection is very important in order to avoid creating a security vulnerability. In the detection phase, relay node locations are divided and clusters are assigned to each UAV using k-means++ clustering. Depending on the number of failed nodes, UAV-NetRest dynamically adjusts the number of UAVs. In this way, network fault detection is performed quickly. The evaluation of the algorithm is based on the detection and the recovery time, the distance traveled by UAVs during the operation, the number of messages transmitted, and the deployment of the relay node.

AI-enabled routing protocols developed for UAV networks are used to detect and prevent threats [47]. For example, a topology predictive and self-adaptive learning-based routing algorithms to adjust the dynamic network topology are used in UAVs. Similarly, the cost of offloading for the user equipment (UE) and the energy efficiency of the UAV system is done by using Markov decision processes in [48]. It presents a Multi-Agent Reinforcement Deep Learning algorithm (MADRL) based on deep learning principles. MADRL aims to optimize power control, resource allocation, and UE association, so that it reduces energy consumption in the system.

The exponential growth of systems that use IoT devices like in IoT devices in health and UAV has led to use machine learning algorithms for better decision making. However, the security threads based on the usage of IoT devices remain [49]. Both IoT usage in health and in UAV have similar security threats. For instance, IoT devices are used in health systems as biomedical sensors, which have similar properties with sensors on UAV. Since sensors may build a wireless network, a secure wireless communication is essential [50]. Moreover, IoT based health and UAV systems may use machine learning algorithms for decision making, which algorithms are prone to data poisoning attacks [51]. Furthermore, systematic poisoning attacks are possible for such critical systems [52].

Understanding and mitigating vulnerabilities require a holistic approach that combines technological innovation, security measures, and legal requirements. Ongoing research and developments are essential to stay ahead of emerging threats in the dynamic landscape of UAV technology.

3.2. *Security attacks*

UAV networks are exposed to different types of attacks. The goal of an attack to an UAV system is to absorb and control network traffic, interrupt the routing function, or inject malicious nodes. Various attacks in the environment of UAV networks have been described in the literature. For example, an overview of cyber attacks that may affect privacy, integrity, and availability of UAV systems is presented in [53]. Security of communication links between ground control station and drones is one of the main research focus [54]. Many articles in the literature focus on identifying security threats so corresponding attacks in drone communications and their corresponding countermeasures. The synthesis of security risks in various systems sets the stage for a discourse on key security challenges that occurred in unmanned systems [55], such as privacy and security challenges in the Internet of Drones (IoD).

A man-in-the-middle attack against a typical UAV used for critical applications is a common example for attacks in UAV systems [56]. GPS spoofing attacks to manipulate the trajectory of an autonomous UAV is another significant attack on UAV systems [57–59]. An attacker can compromise protocols using various attack strategies. Attacks and countermeasures for security protocols are highlighted, and a schematic diagram of possible attacks is provided in [60]. In general, when we evaluate all researches together in the literature, main attacks carried out on UAVs are GPS Spoofing Attacks, False Data Injection, Jamming Attacks, Routing Attacks, Eavesdropping, Side-Channel Attacks, Code Modification, Code Injection, Packet Sniffing, Replay Attacks.

3.2.1. *False data injection attacks*

False data injection is a technique where an unauthorized person sends copied data to the UAV to take it control as shown in Fig. 5. UAVs are exposed to malicious data and they may not distinguish between real and malicious data. To perform a false data injection, attackers typically take control of UAVs by injecting fake data into the UAVs' sensors using attackers' sensors.

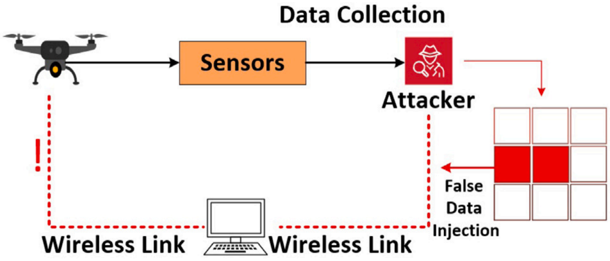


Fig. 5. False data injection attack on UAVs.

Countermeasures. There are many countermeasures depending on the application of the UAV. In [61], a new algorithm is introduced to detect a data injection attack of a UAV bug. An adaptive neural network is used to detect errors injected into the sensors of the UAV. An integrated Kalman filter is used for in-line adjustment of the neural network weights. This inline setting makes intrusion detection faster and more accurate. Samy et al. introduce an NN-based system model for the fault detection process [62]. In this work, the system identification is achieved through the offline learning process, so the fault detection process does not need model data during the operation. A NN-based flaw detection design for failures of actuators on a satellite is presented in [63]. Shen et al. presented a neural network-based fault detection technique that takes into account the latency between a fault detection and a fault compliance [64].

3.2.2. *Routing attacks*

An adversary attacks existing routing protocols to reduce the performance of an UAVs network or change its topology [35]. An attacker can degrade the performance of the UAV network by corrupting the routing algorithm or launching a DOS attack. An UAVs network's topology may be altered by adding non-existent nodes to routing tables, creating a fake route connection. In the literature, routing attacks may be explained within a controlling network traffic, interrupting routing functions, or injects malicious nodes. However, routing protocols may be exposed to different types of attacks. In literature, path discovery attacks, path maintenance attacks, and data transmission phase attacks based on their basic routing functions are analyzed as routing attacks [65]. [35] contains attacks on the routing protocol in a classified manner.

Countermeasures. In these types of attacks, cryptographic methods and intrusion detection systems are used by an UAV on a communication line. For example, the Security Sensitive Temporary Routing (SAR) protocol uses security metrics in Route Request Packet (RREQ) [66]. It also uses a shared secret to generate a symmetric encryption key. A location-based routing solution to provide authentication and privacy is possible [67]. An anomaly-based detection mechanism that learns from the statistical analysis of different RREQ packet rates and calculates the threshold instantly is proposed in [68].

3.2.3. *Eavesdropping attacks*

In an eavesdropping attack, the attacker passively or actively intercepts network communications to access secret information on an UAV network, such as node ID numbers, routing updates, or application-sensitive data. An attacker can use this private information to compromise network nodes, alter routing, or reduce application performance. Active eavesdropping aims to attack the main channel by reducing the channel capacity. Active spy transmits interference noise and simultaneously and independently captures spy signals [69]. Some eavesdroppers relay on the interference of a legitimate receiver, while others intercept the nosy.

Countermeasures. Encryption is the standard defense against eavesdropping attacks. Due to limited processing power of some UAV systems in wireless communications, they cannot efficiently handle the standard encryption methods used in typical wired networks [70].

3.2.4. Side channel attacks

A side-channel attack refers to the disclosure of useful information about the internal execution of a system, either with transmitted data or via alternative routes. This type of attack obtains information indirectly by exploiting information leakage. Examples of side-channel attacks on UAVs include acquisition and analysis of meteorological information, power consumption, electromagnetic dissipation, acoustic signal analysis, and residual data. Defense against side-channel attacks includes the use of asynchronous processing units and mechanisms that help reduce electromagnetic emissions. The temporal information of the Unmanned Autonomy Systems communication channel is strongly associated with sensitive internal states, such as the number of UAVs.

Countermeasures. Two possible solutions can be used to counter side-channel attacks [71]. Unmanned Autonomy Systems may randomize the temporal information of packets by generating redundant packets within random time intervals. While this solution may eliminate the loss of confidential information from the communication channel, it may in turn reveal the existence of the Unmanned Autonomy Systems communication channel, which may allow other attacks, such as a signal interference. On the other hand, Unmanned Autonomy Systems can use traffic transformation to statistically make Unmanned Autonomy Systems traffic patterns indistinguishable from the traffic patterns of popular network applications, such as web clients or messengers [72]. Additionally, side-channel resistant implementations on devices in an UAV may help to have better security. For instance, Edwards Curve Digital Signature Algorithm (EDDSA) based on the Ed448 targeting the ARM Cortex-M4-based STM32F407VG micro-controller on IoT devices in an UAV system may help to prevent side channel attacks on UAV systems [73]. Since current systems have been expected to be broken by the advent of quantum computing, post quantum solutions are expected to be implemented for critical systems, such as cryptographic accelerators for digital signatures that provide high performance Ed25519 architecture [74]. Integrated countermeasures to high performance post quantum computing and their optimized version are expected to be emerging implementations on IoT devices for UAV systems [75].

3.2.5. Code injection attacks

After gaining access to the UAV networks and control units, an attacker can inject malicious code into the control units. Malicious payloads such as viruses, Trojan, and spyware may also infiltrate anti-malware software with this approach. UAV controllers may also inject code when one or more UAV components or subsystems are incompatible, in hopes of improving their vehicle's performance or misleading regulatory examination.

Countermeasures. The defense against code injection attacks may be applying an intrusion detection system. In this case, the access control system should only grants permissions to authorized personnel. Additionally, UAVs may not include vehicle owners under certain circumstances that case need to be carefully considered [76].

3.2.6. Replay attacks

A replay attack may destroy cyberphysical systems, including UAVs, even without knowing the external structure. The replay attack is basically shown in Fig. 6. Replay attacks are the simplest attack that can be implemented for malicious purposes. An attacker records and covers transmitted data from a network perspective, compromising closed-loop capabilities of cyber–physical systems and degrading system performance.

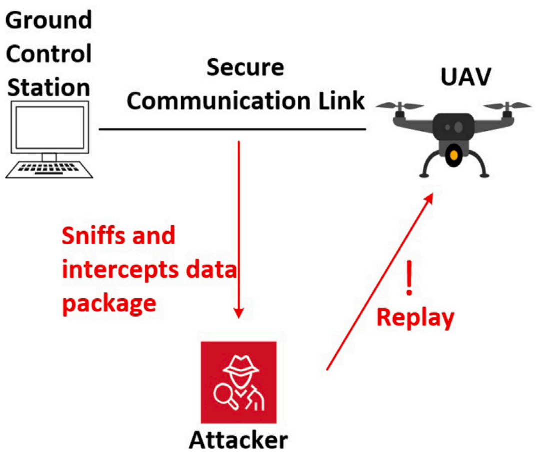


Fig. 6. A replay attack on an UAV.

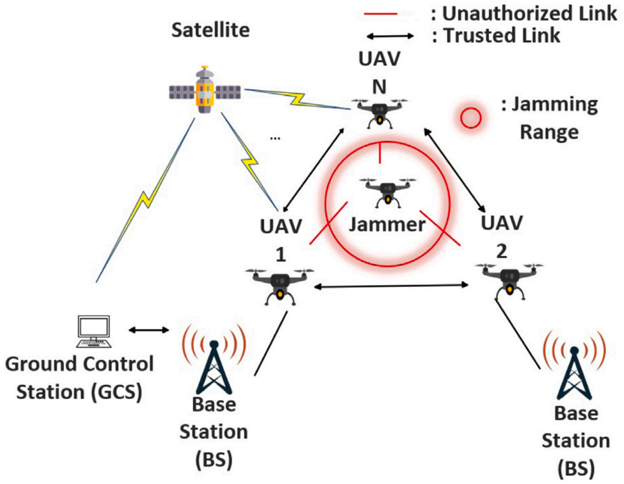


Fig. 7. A jamming attack on an UAV.

Countermeasures. A countermeasure against replay attacks is a recording horizon control law to address the attack on the communication network between the controller and the actuators [77]. Under bandwidth constraints, a secure fusion prediction scheme for cyber–physical systems against replay attacks has been proposed in [78]. Based on this detection measure, a stochastic play approach has been developed to reduce the loss of the control performance [79].

3.2.7. Packet sniffing attacks

A packet sniffer may intercept and log traffic transmitted over a communication link. This type of attack is derived from a tool that is commonly used to diagnose network-related problems. An attacker can use a packet sniffer to spy on unencrypted data in packets to gather information.

Countermeasures. Possible defenses against packet monitoring include the application of encryption techniques to protect the confidentiality of packets in transit, as well as distribution techniques to secure and authenticate sent and received communication signals [76].

3.2.8. Jamming attacks

In UAV systems, jamming attacks are defined as electromagnetic energy emitted in a deliberate direction to a communication system to interrupt or prevent the transmission of the signal [80] as shown in Fig. 7.

A jammer may be anything, such as from being a single specially equipped transmitter to all jamming stations. The goal of broadcast jammers is to reduce availability in a secure system. A comparative analysis of the various jamming techniques is discussed in [81].

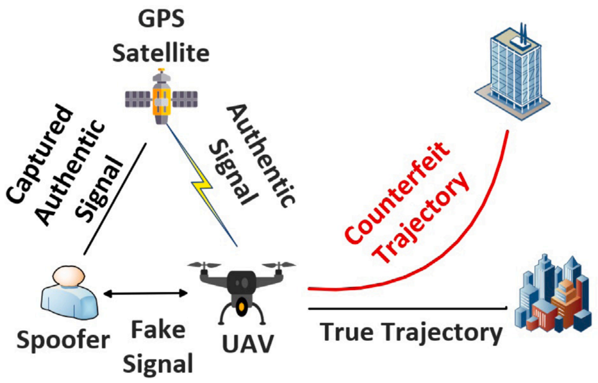


Fig. 8. A GPS spoofing attack on an UAV.

Countermeasures. Classical countermeasures are block jamming attacks, moving away from the disrupted field hopping frequency, distributed attack detection, distributed secure predictions, and statistical approaches in wireless sensor networks [82,83]. Especially in VANETs, classical countermeasure methods are not always applicable due to high mobility and the wide network topology.

3.2.9. GPS spoofing attacks

The Global Positioning System (GPS) is typically used to provide UAV navigation information and time information. However, GPS is subject to many deliberate threats. For example, an attacker can mislead a GPS receiver with false GPS signals as shown in Fig. 8. First, a reliable navigation system is required for new autonomous systems based on the standalone UAV. GPS sensors are the most widely used navigation sensors for high-performance flights in UAV navigation systems. In military applications, GPS signals are encrypted to prevent unauthorized use and spoofing. However, the current civil GPS signal is transparent and easily accessible worldwide, making GPS-guided civilian infrastructures vulnerable to various frauds or broadcast jammers [84–86]. There are also experimental studies to examine the GPS signal generation process and effects of spoofing signals on UAVs [87].

Countermeasures. The recommended countermeasure is to prevent spoofing attacks on civilian UAVs is to encrypt GPS signals. However, this is quite costly since it requires significant upgrading of the infrastructure. There are various anti-spoofing techniques to detect a false signal. The control of WLAN interaction points, a cross-check detection control, an activation of an alarm system when the ratio between the signal strength and the noise exceed a certain threshold is one of them [88].

There are more advanced attacks on GPS that are relatively difficult to detect. For example, a receiver-based spoof system is more complex, which consists of a GPS receiver that is combined with a phishing transmitter. The generated signal is synchronized with real GPS signals [31]. Research is carried out to confirm whether the error between the UAV model estimator and the real signal is fake or not [89]. Threshold-based methods often increase the complexity of software and hardware. Moreover, the detection of a spoofed signal is not guaranteed when there is a very sophisticated spoof. Therefore, innovative and effective approaches are needed to secure GPS services in UAVs.

4. Traditional defenses for UAVs

UAVs are vulnerable to a variety of attacks, including data interception, malicious data injection, jamming, and spoofing. These attacks may lead to a loss of control over the UAV, data theft, or damage to the UAV itself. Various defense techniques have been developed to prevent these attacks, including encryption methods, authentication mechanisms, and intrusion detection systems. This section provides an overview of these techniques against security attacks on UAVs as shown in Fig. 9.

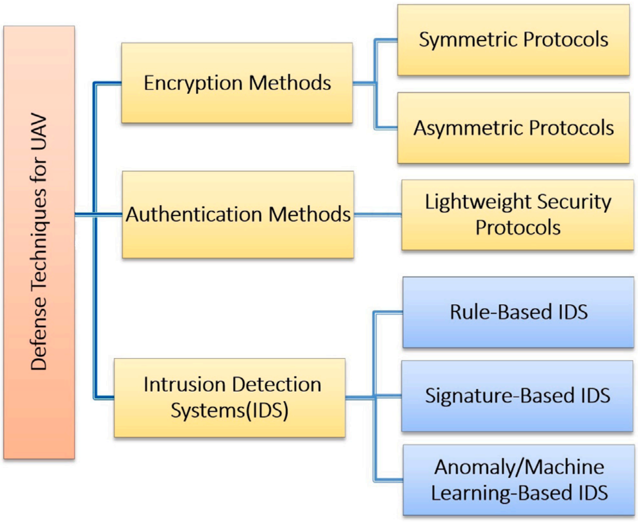

Defense Techniques for UAV
Encryption Methods
Symmetric Protocols
Asymmetric Protocols
Authentication Methods
Lightweight Security Protocols
Intrusion Detection Systems(IDS)
Rule-Based IDS
Signature-Based IDS
Anomaly/Machine Learning-Based IDS

Fig. 9. A summary of defense techniques for UAV security.

Encryption and authentication. Basic defense techniques against security attacks on UAVs are encryption and authentication. Encryption ensures that data is transmitted securely by encoding it in a way that only authorized users can decipher. Authentication ensures that only authorized users are allowed access to the UAV, its control systems, and its data. Cryptographic techniques are used to guarantee confidentiality, integrity, and availability of the system. Specifically, symmetric cryptographic protocols are employed to safeguard sensitive information, such as text, audio, video, and images. In these protocols, both the sender and receiver have a shared key for both encryption and decryption purposes.

Asymmetric security protocols use a pair of distinct keys, namely a public key and a private key, for encryption and decryption of data by the sender and receiver, respectively. Unlike symmetric protocols, the confidentiality of the public key is not a critical issue, since encrypted data cannot be decrypted using the same key. Hence, a private key is always necessary for data decryption.

A lightweight authentication protocol is a security protocol designed to provide authentication of data or entities while minimizing the overhead on the system resources, such as memory and computational power. These protocols are particularly useful for resource-constrained devices such as UAVs, where traditional cryptographic protocols may be too heavy to be implemented. The goal of a lightweight authentication protocol is to provide sufficient security with minimal system resources. These techniques are commonly used to protect data transmitted between the UAV and the ground control station [90].

Encryption techniques in flying UAVs are also used to prevent attacks that cause hardware failures [60]. Thus, an enemy is prevented from obtaining confidential data. An authenticated encryption mechanism is used against eavesdropping attacks in similar cases [91]. Additionally, there are secure communication schemes against replay attacks [60,92].

Secure transmitter systems against hijacking, eavesdropping, and Distributed Denial-of-Service (DDoS) attacks are implemented using encryption methods. A System Development Life Cycle (SDLC) is used with a prototyping approach, which is implemented using three prototypes in [93]. The proposed scheme employs an Android application for a web platform and Advanced Encryption Standard (AES) algorithm for secure transmissions. Security tests have shown that the proposed scheme may resist attacks with a high success rate. The User Acceptance test has also confirmed that the application used for the secure transmission meets the user's requirements. In [94] proposed a method to verify the authenticity of data, which ensures that data were sent by the original GCS and not by any unauthorized source. The proposed method uses asymmetric protocols to protect the integrity of data transmitted between different sensors or devices.

Hash algorithms are used in various parts of protocols to provide security for UAVs. For example, the verification of signatures after data is received by the recipient uses hash algorithms as in [95]. The UAV performs the verification process to ensure the confidentiality of data before doing an action. A 164-bit hash is generated using the SHA-1 algorithm on the sender side for verification of data. The generated hash is then encrypted using a public key and then sent to the recipient. The receiver uses a private key to decrypt the encrypted hash value. Then, it calculates the hash value from the data from the original message. The two checksums are then compared to verify the authenticity of the message.

Elliptic Curve Encryption is one of the significant cryptographic algorithm that is used in many protocols. In [96] proposed a lightweight recognition mode for authentication using Elliptic Curve Encryption (ECC) against UAV network attacks. The proposed approach is designed to protect against different types of security attack, such as message eavesdropping, fake identity, and replay attacks on a UAV network. The approach uses several security measures, such as an ECC digital certificate for UAV identification credentials, ECDSA for UAV identification and signature verification, and the Elliptic Curve Diffie Hellman (ECDH) algorithm for session key negotiation during drone communication.

Lightweight security protocols are important for security of UAVs because the protocols consume relatively small amount of energy. In [97] proposed a lightweight security protocol for IoD networks called Temporary Credential Based Anonymous Lightweight Authentication Scheme (TCALAS) that is limited to a single flight zone. This uses a threat model and an authentication. The protocol also includes a Ground Station Server (GSS), remote drones, a mobile device, and a control room. An upgraded version of the protocol is introduced in [98], which may provide security against various attacks and be implemented in multiple flight zones. The network model of the proposed scheme consists of a GSS, a control room, drones with related flight zones, and a drone user. The proposed scheme provides security features, such as anonymity and non-traceability of the user, mutual authentication, and robustness. This version of the protocol has also been shown to be faster to complete the entire authentication process.

Traditional security mechanisms use symmetric and asymmetric cryptography to implement security services. One of the significant threats to these services is the advent of quantum computing that requires post-quantum algorithms to be implemented and integrated to security services. UAV systems already use security services. Therefore, efficient implementations of post-quantum algorithms with different technologies, such as FPGA, are needed to protect UAV systems [99]. For example, high-speed NTT-based polynomial accelerators for post-quantum cryptography are required [100]. Another implementation may be on ARM Cortex-M4 [101]. A more detailed analysis of post-quantum computing is presented in [102], which covers quantum computing and post quantum computing for blockchain. All of these security services are expected to be implemented for UAV systems.

Intrusion detection systems (IDSs). (IDSs) are another technique used to defend against security attacks on UAVs. These systems monitor the UAV's data traffic and alert the operator if any unauthorized or malicious activity is detected. IDSs may also monitor the UAV's physical environment, such as detecting any attempt to tamper the UAV's hardware or software. In this paper, intrusion detection systems are examined in three categories.

1. **Rule based IDS:** Rule-based intrusion detection is performed. In this category, rules following the expected behavior of the UAV system are applied on UAVs to perform special tasks [103]. Rule based IDS sets specific rules and policies that define network behavior. Any deviation from these rules is detected as a potential intrusion.
2. **Signature based IDS:** This category involves the comparing network traffic with a known signature database or known attack patterns. If a match is found, an alarm is triggered and the intrusion is detected. A Bayesian-based technique for detecting insiders and removing malicious nodes from the network is presented in [104]. The technique includes activating an intrusion detection mode for various nodes and calculating the false behavior rate (MR) for nearby UAVs. If the threshold is lower than MR, the intrusion detection system starts monitoring nearby nodes and detects intrusions. The intrusion removal system determines the MR of a node, and if its threshold is less than the MR, the node is declared rogue and it is removed from the network.
3. **Anomaly-based/Machine Learning-based IDS:** This category involves detecting known and unknown attacks based on learning or filtering mechanisms. The working principle of Anomaly-based/Machine Learning-based IDS is identifying unusual or anomalous behavior in the network. It uses statistical analysis or machine learning algorithms to distinguish the normal behavior and detects any deviation from the normal behavior. Machine learning (ML) algorithms have two phases, namely training and testing. The model is trained to predict future events using training data in training phase. The testing phase uses various strategies to measure the model's accuracy. A summary of ML-Based Identities for UAV Security is given in Fig. 10.

Each IDS method has advantages and disadvantages. For example, signature-based IDS is effective for detecting known attacks but it is weak against new or unknown attacks that change their patterns frequently. If one bit changes, the signature changes, too. Anomaly-based IDS may detect unknown attacks, but they may also suffer from false positives and false negatives. A more efficient IDS is to use a hybrid detection approach that combines two or more approaches for the accurate detection of unknown attacks [105]. A hybrid IDS can strike a balance between the two and is often the preferred approach for UAV security.

Recently, ML-based IDS becomes popular. For example, a machine learning-based IDS was proposed against GPS Spoofing attacks in [106]. The recommended method includes the one-class support vector machine and ML autoencoder algorithms. A method for detecting eavesdropping attacks on UAV communication is presented in [107]. The method uses K-mean clustering and support vector machine (SVM) algorithms that may learn from existing data and make decisions for future samples. The technique consists of two phases. In the first phase, both parties send signals to the UAV. In the second phase, the UAV transmits the same signals to a third party to detect any deviation. A dataset is created and classified using machine learning algorithms to identify potential attacks.

ML-based IDS is used against signal spoofing and jamming attacks. In [108], Self-Teaching (STL) with a multi-class SVM is used to maintain a high true positive rate for IDS. It uses Deep-Q Network, a deep reinforcement learning algorithm for the drone dynamic route learning as a self-healing method in the IDS recovery phase. The proposed solution collects data from UAV components, such as flight logs and reading data from sensors. However, a real-world application of UAVs is difficult due to the limited power and a lack of enough computational resources.

A convolutional neural network (CNN) method is used to detect jamming signals in [109]. The UAV algorithm takes into account the weights and values of GCS, selects a relay power element based on Bit Error Rate (BER), and uses the boost and randomly selected values to send a message. While the algorithm may protect communications and jamming attacks, randomly selected relay power may increase the error rate, making it expensive.

Several machine learning-based security frameworks have been introduced in the literature to address a variety of security issues, such as

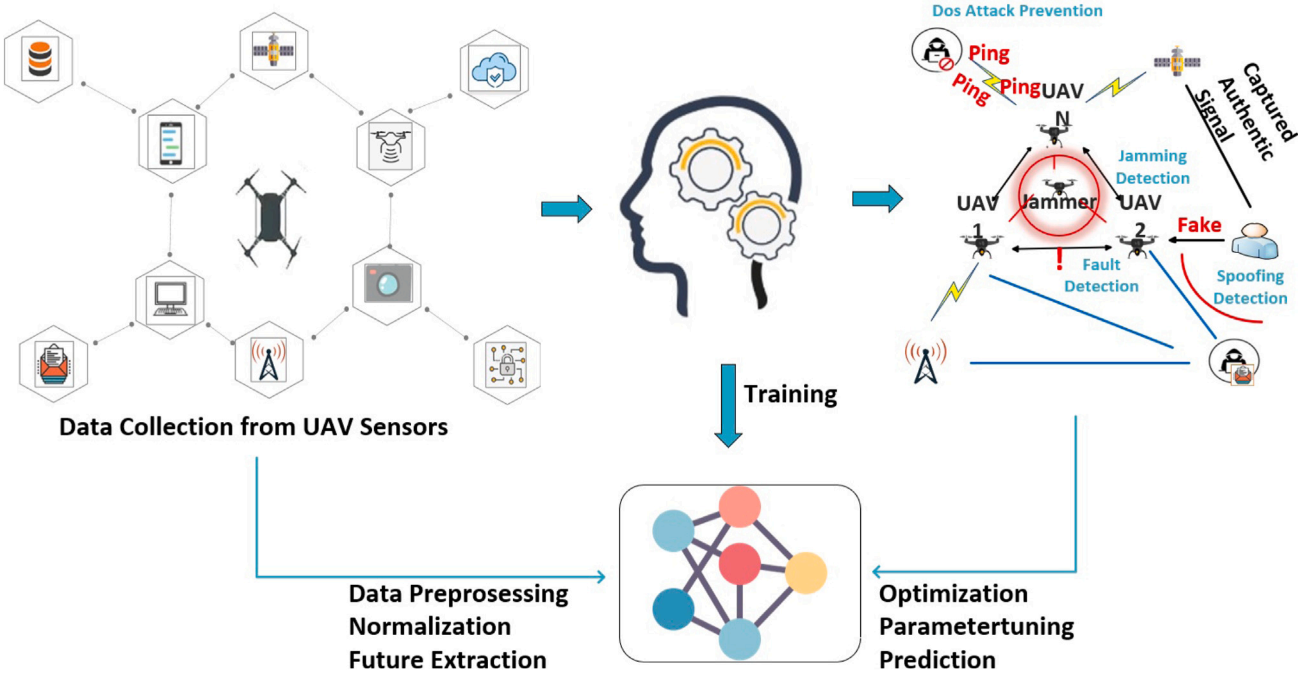


Fig. 10. A summary of ML-based IDS for UAV security.

malicious drone detection and DoS attack prevention for UAVs [110]. Some recent research has shown that federated learning techniques may be more effective than traditional machine learning algorithms. For example, there are reports on the development of radio frequency-based UAV authentication models using IoT networks [111].

A Lightly Distributed Detection Scheme aims to detect and mitigate flooding attacks in the Internet of Drones (IoD) environment that is explained in [112]. The scheme uses the concept of automatic count reports, where each drone counts the number of packages it sends in a given time interval and shares that report with other drones during contacts. Receiver drones store reports and send them to nearby ground stations to check consistency and detect flood attacks.

Specific attacks may require specific countermeasures. Protection against specific attacks is achieved through path planning, with illustrative examples in [129]. An airway-aware protection mechanism (IoD-JAPM) against jamming attacks (JA) on the Internet of Drones (IoD) is presented, including an analysis of airway availability and potential modifications on drone path planning. Specifically, IoD-JAPM effectively uses jamming attacks on all drones in the presence of a jammer. The approach contains three key stages, which are airway analysis, risk region (HR) discovery, and route planning generation.

A hybrid IDS model integrating spectral traffic analysis and a robust controller/observer for anomaly estimation within UAV networks is presented in [113]. The IDS is created against DDOS attacks. The hybrid method takes into account, in its initial phase, a statistical representation of the network's traffic. By analyzing the newly created signatures, anomalies may be determined with an appropriate model to accurately estimate the anomalous traffic.

A summary of the state of the art on attacks and countermeasures is given in Table 1. We observe that each attack focuses on a different target on UAVs and violates different security requirements. Therefore, it is important to note that a single defense technique is not sufficient to protect UAVs from security attacks. A combination of these techniques must be applied to provide comprehensive security to UAVs. Moreover, the power consumption and battery usage must be evaluated according to the target of the attack to determine potential countermeasures. In addition, the defense techniques used should be regularly updated and tested to ensure that they are effective against the latest security attacks.

5. Reinforcement learning and UAV security

This section contains a comprehensive review of the state-of-the-art Deep Reinforcement Learning (DRL) based techniques in the field of security in UAVs. In addition, basics of RL techniques are given to make better understandable how RL techniques are used in UAV security.

5.1. Reinforcement learning

Reinforcement learning describes machine learning algorithms that try to find optimal behavior in an environment that maximizes the accumulative reward. A common structure of reinforcement learning for UAVs is shown in Fig. 11. The decision-making process controls the environment that is often considered a Markov Decision Process (MDP) [11]. An MDP consists of a set of states S, which may be finite, infinite, or continuous, a set of executable actions, which can be discrete or continuous, a transition probability, which describes the dynamics of an agent interacting with its environment, and a reward function at a given initial, final state after taking a certain action.

Classic RL algorithms have limitations when large-scale problems need to be solved with RL algorithms. For example, many classic RL algorithms require to maintain a value look-up table for each state or a state–action pair [11]. Thanks to the recent advances in computational power, Deep Learning (DL) overcome this limitation. Most of the time, RL algorithms use Temporal Difference Learning and Q-learning for Actor-Critic models.

Temporal Difference (TD) learning consists of a transition model and a reward function with policy evaluation [130]. The state value function of MDP, $V_\pi(s)$, is calculated as follows.

$$V_\pi(s) = E[\sum_{t=0}^{\infty} \gamma^t R(s_t)|s_0 = s, \pi] \tag{1}$$

In the standard dynamic programming approach, $V_\pi(s)$ is computed as follows.

$$\forall s: \quad V(s) \leftarrow R(s) + \gamma \sum_{s'} P(s'|s,a)V(s') \tag{2}$$

A stochastic version of the state function is as follows.

$$\text{pick some state } s: \quad \text{sample } s' \sim P(s'|s,\pi(s)) \tag{3}$$

Table 1
Attacks and countermeasures for UAV security.

Ref.	Vulnerability	Attack	Target	Defense	Security requirement
[92], [60]	Wireless Link	DOS	UAV Network	Authenticated Encryption	Confidentiality, Authentication
[93]	Wireless Link	Hijacking, Eavesdropping, DDOS	IOD Communications	AES Algorithm	Availability, Authentication
[94]	Automatic Dependent Surveillance Broadcast (ADS-B)	Jamming, False Injection	Air Transportation System	Asymmetric Encryption	Authentication, Integrity
[95]	ADS-B	Data Spoofing	Air Transportation System	ECC, ECDSA Algorithm	Confidentiality, Authentication
[96]	Wireless Link	Eavesdropping, Replay Attacks, Fake Identity	UAV Network	ECDH Authentication using ECC	Confidentiality , Authentication
[97]	Wireless Link	Identity Attack, GSS or Remote Drone Impersonation Attack, Replay Attack	IOD Network	TCALAS	Authentication
[98]	Public Communication Link, Power Analysis Leakage, Public servers and users	Identity Guessing Attack, Impersonation Attacks, Replay attack, Man in the Middle Attack, Remote Drone Capture Attack	iTCALAS	IOD Network	Anonymity, Non-traceability, Authentication, Robustness
[106]	A wide variety of sensors, broadcasting fake GPS signals	GPS Spoofing	UAV Sensors	One Class Vector Machine, Autoencoder (ML based IDS)	Confidentiality Integrity, Authenticity, Availability, and Accuracy of GPS data
[107]	Wireless Link	Eavesdropping	UAV communication	K-means, support vector machine (SVM) algorithms (ML based IDS)	Authenticity
[108]	Broadcasting fake signals	Jamming, Signal Spoofing	UAV communication	Self-Teaching (STL) with a multi-class SVM (ML based IDS)	Accuracy, Sensitivity and Specificity, Confidentiality
[110]	Wireless Link	DOS	UAV communication	SVM, CNN	Identify
[111]	Radio Frequency (RF) vulnerability	Privacy-Preserving, Data security	IOT Network	Federated Learning	Authentication
[112]	Storage space, Packets Size	Flooding Attacks	IOD Environments	Lightly Distributed Detection	Identify
[113]	Changes of network topology, mobility	DDOS Attacks	UAV ad-hoc communication	Hybrid IDS Model	Availability
[114], [115], [116]	Radio Frequency Vulnerability	Jamming	Cognitive Radio Network	DQN	Availability
[116]	Large-scale dynamic network, high mobility	Jamming	VANET	PHC	Availability
[117]	Dynamic Radio Environments, wireless link	GPS Spoofing	Wireless Network	Q Learning, Dyna-Q	Authentication, Identify
[118]	Wireless Link	Jamming	IOT Networks	DDPG, TD3	Robustness, Availability
[119]	Wireless Connectivity	Jamming	Multi-UAV Cellular Network	TD Learning	Availability
[120]	Radio Frequency Vulnerability	Eavesdropping, Jamming	Multi-UAV Cellular Network	MDP, Double DQN	Stability, Availability
[121]	Offloading System	Jamming	UAV Communication Network	DDPG	Availability, Confidentiality
[122] [123]	Wireless Link	Eavesdropping	UAV Communication with RIS	TTD3	Confidentiality
[124]	Wireless Link	Eavesdropping	IRS-assisted UAV covert communication system,	TAP-DDQN	Confidentiality
[125]	High Mobility, Large Scale Network Topology	Jamming	VANET	PHC	Availability, Confidentiality
[126]	Broadcasting Incorrect GPS Spoofing	GPS Spoofing	Authentic Signal Received from GPS Satellite	Game-Theoretic Security Mechanism	Reliability
[127]	Energy consumption, computation capacity and network delay	DOS, Offloading attacks	Unmanned Aerial Vehicle Edge Computing (UEC) network	Stackelberg Game-Theoretic Security Mechanism	Availability
[128]	Openness of Networks	DOS, Unauthorized Access Attacks (R2l, U2r), Probe Attack	UAV Network (NSL-KDD Cup dataset)	DRL-BWO Algorithm	Availability

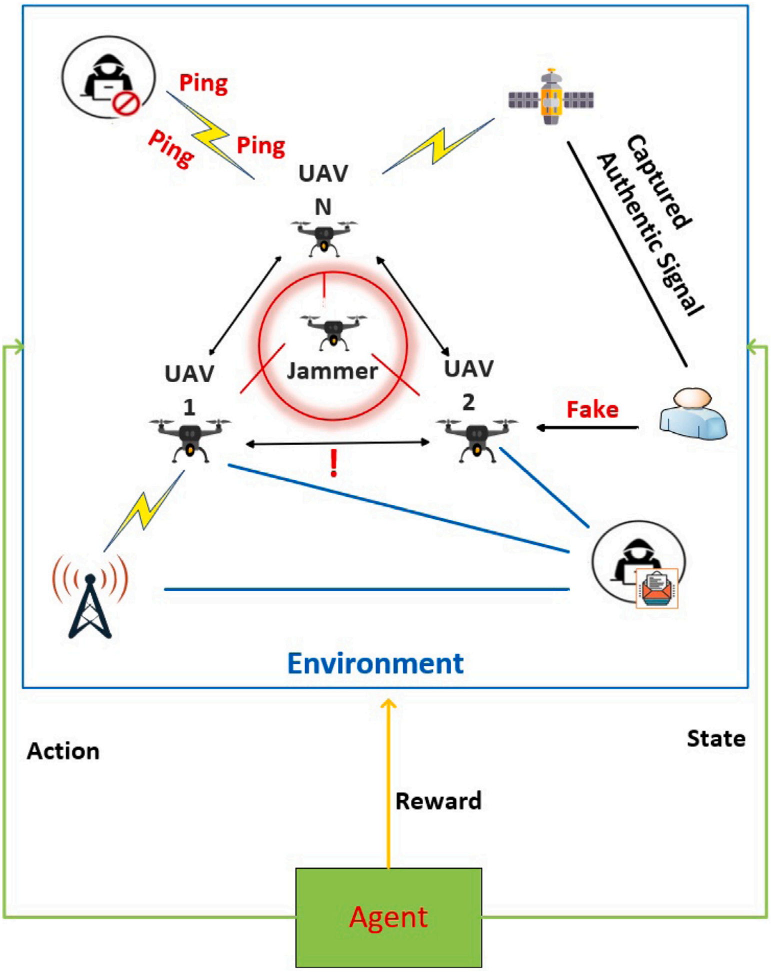


Fig. 11. General structure of reinforcement learning for UAV security.

$$V_\pi(s) \leftarrow \alpha(R(s)+\gamma V'_\pi(s)) + (1-\alpha)V_\pi(s) \tag{4}$$

where α is the step size. A possible step size is $\alpha_k = \frac{1}{k}$ for the k'th time for an update.

Convergence is the stochastic version of policy evaluation that converges to the true value function under certain assumptions. The following assumptions are sufficient conditions.

- every state is often visited infinitely.
- α satisfies $\sum_{k=0}^{\infty} \alpha_k = \infty$; $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$.

In practice, a reason to use TD for policy evaluation could be that we do not have the transition model available. The samples are then generated by executing the policy and by performing stochastic value function updates according to the states that are visited.

TD only considers policy *evaluation.* If we are interested in finding a near-optimal policy, we can use TD as the policy evaluation step in *policy iteration* as follows.

Iterate:

- Run TD to perform policy evaluation, which gives us $V_\pi(s)$, $\forall s$.
- Pick a new policy π, such that $\pi(s) = \arg\max\ [R(s) + \gamma \sum_{s'} P(s'|s,a)V(s')]$

Iterate:

$$\forall s, a : Q(s,a) \leftarrow R(s) + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q(s',a') \tag{5}$$

Similarly to TD learning, we can run a stochastic version instead for k = 0, 1, 2,…, as follows.

$$Q(s,a) \leftarrow (1-\alpha_k)Q(s,a) + \alpha_k(R(s) + \gamma \sum_{s'} \max_{a'} Q(s',a')) \tag{6}$$

The stochastic version does not require a priori knowledge of the transition probability distribution P or the reward function R. It is sufficient to have access to a trace. The stochastic version may converge to the true Q function under the same assumptions as in TD.

- every state is often visited infinitely.
- α satisfies $\sum_{k=0}^{\infty} \alpha_k = \infty$; $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$.

In practice, one can ensure that all reachable states are visited infinitely often by using an ϵ greedy policy;

with probability ϵ : choose an action at random
with probability $1-\epsilon$: $a = \arg\max_a Q(s,a)$

Another popular way to choose the actions is computed as follows.

$$\text{pick action } a \text{ with probability} \frac{\exp(-Q(s,a)/T)}{\sum_b \exp(-Q(s,b)/T)} \tag{7}$$

In Eq. (7), T is the temperature and it determines how greedily the actions are being chosen. A natural choice decreases the temperature over time. Note that the temperature T and factor α need not be the same for all states. For instance, one could have these variables that depends on how often the current state s has been visited.

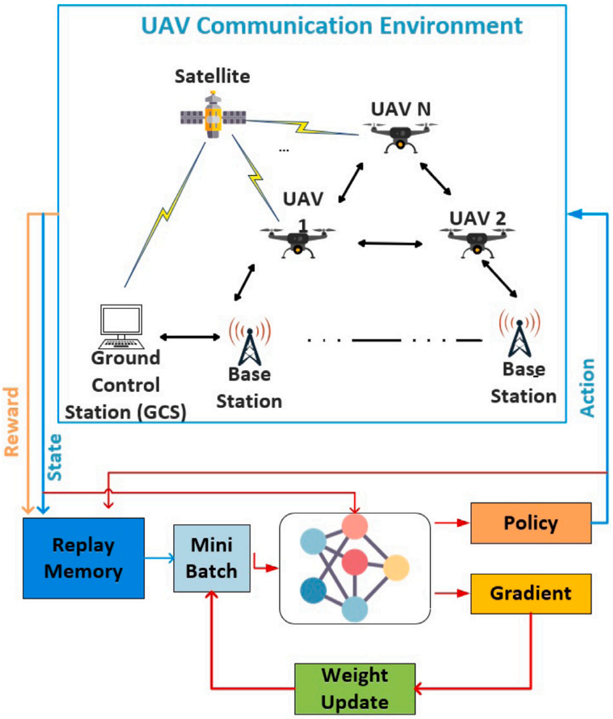


Fig. 12. An example for DQN.

Q-learning. is one of the commonly used algorithms to solve the MDP problem. The actions are obtained for every state that is based on an action-value function. *Q(s, a)* is defined with the value of the state (*s*) and action (*a*) pair for attacks in communication networks. *R(s)* is the reward function of the current state. *P(s'|s, a)* is defined as the probability of the transition from the actual state–action pair to the next attacking state in a threat condition. *V(s')* is the value of the next state. *Q(s', a')* can be defined as the value of the next state (*s'*) and next action (*a'*).

$$Q(s,a) = R(s) + \gamma \sum_{s'} P(s'|s,a)V(s') \tag{8}$$

We can run dynamic programming, such as value iteration, by performing the Bellman back-ups in terms of Q function as follows [11].

$$Q(s,a) \leftarrow R(s) + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q(s',a') \tag{9}$$

Deep Q networks. When states are discrete, Q-function can be easily formulated as a table. This formulation becomes harder when the number of states increases, and may be impossible when the states are continuous. In such a case, the Q function is formulated as a parameterized function of the states, actions pairs *Q(s; a;w)*. The solution is required to find the best setting for the parameter *w*. Using this formulation, it is possible to approximate Q-function by using a Deep Neural Network (DNN). In this setup, the problem will be an optimization problem. DNN may minimize the Mean Square Error (MSE) of Q-values using Gradient-based methods Stochastic Gradient Descent (SGD) [131]. Fig. 12 shows an example of a Q-learning mechanism from a neural network perspective.

5.2. Reinforcement learning for UAV security

RL-based solutions in UAV security cover a wide range of applications, including strategies for the protection of cyber–physical systems and communication technologies, autonomous attack detection methodologies, and solutions that are based on game theory principles.

In the literature, attacks have targets on all components of the UAV network infrastructure. Therefore, there is a need for environment independent methods to detect various attacks quickly. Deep Reinforcement Learning methods have huge potential to detect attacks on UAV systems in a faster manner. DRL may also provide the most convenient countermeasures for UAVs by ensuring a self-adaptation that is independent from the environment and resists to many attacks. A summary of deep reinforcement learning applications in communication networks is presented in [34]. In this paper, we discuss the recommended reinforcement learning applications for detecting jamming and spoofing attacks targeting communication in UAVs.

In modern broadcast jammer prevention techniques, an attack-free self-learning protection system protects the communication with deep reinforcement learning. In the literature, broadcast jamming attacks have been discussed in different communication environments, such as Cognitive radio network (CRN), WSN, 5G, GPS, VANET, and MANET in UAV. Deep reinforcement learning techniques vary according to the area of usage in UAV systems. For example, DRL effectively decides power allocation, relay or not, channel selection, user selection, and resource allocation. Moreover, different reinforcement learning practices may adapt very quickly in natural environments.

Broadcast-busting attacks may have disruptive consequences for UAVs. Although there are solutions to prevent broadcast-busting attacks, they are inefficient in most cases [114]. DL methods may help to create efficient solutions, such as a dynamic game computational radio channel model against jamming attacks in Deep-Q-Network (DQN) communication. DQN can be used to control the power in a jamming attack targeting the communication [115]. In this case, devices decide to transmit energy by performing the selected action. The Signal to Interference Noise Ratio (SINR) value is measured in a time frame and it is calculated at the end of that time frame. The news channel is tuned in Universal Software Radio Peripherals (USRP). A reactive jammer calculates the utility to select the blocking power that is based on the final transmitted power of the transmitter. At each time slot, the disruptor observes the final SINR and then selects the parasitic power that can be exploited with a greedy strategy. The transmitter selects and adjusts the transmit power in each time slot and transmits data packets to the radio station. It selects the transmission power according to the DQN algorithm to improve the SINR of the transmitter and reduce the energy loss. DQN-based power control strategy improves communication efficiency compared to the Q-learning-based strategy.

Game theory and Q-learning are used to avoid smart jammers, who are targeting UAV communications in VANETs. For example, a game model against jamming is proposed in [116]. In the game, the UAV decides whether to transmit the message while choosing the smart jamming strength. The transmission decision of the UAV depends on the channel quality and Bit Error Rate (BER). An initial anti-disruptive transfer strategy, which is called Policy Hill Climbing Algorithm (PHC) algorithm, is used to obtain the optimum transfer strategy without knowing the VANET model and the compression model [132]. This model reduces the BER in VANET and increases the utility of the UAV compared to the Q-learning-based transmission strategy [125].

UAVs use 5G networks for their communications. Jamming prevention is critical in UAV communications with 5G networks. In this case, it is recommended to move the transmitter away from the faulty area to prevent broadcast jamming, which may interrupt the communication or simply creates communication problems. If 5G networks are used, it is recommended for the UAV to use a fast DQN, learning techniques, deep learning, deep reinforcement learning, and transfer learning together to decide the transmission power control mechanism without knowing the broadcast jammer and the 5G network model [133].

A game-based defense mechanism against GPS spoofing attacks for civilian UAVs is possible as in [126]. For example, a zero sum game using the Stackelberg methodology to represent strategic interactions between security agents deployed in UAVs and potential attackers targeting the Unmanned Aerial Vehicle Edge Computing (UEC) network is presented in [127]. This non-cooperative game requires security agents to protect UAVs and all communication links, which include U2U communication between UAVs and U2I communication between UAVs and infrastructure. It is possible to detect spoofing attacks faster and in more adaptive manner by using Q-learning [117].

Black Widow Optimization (BWO) algorithm, which is specially designed for UAV networks, is used with a DRL technique for optimization purpose in [128]. This approach aims to improve the intrusion detection performance of UAV network attacks. BWO algorithm is used for parameter optimization of the DRL technique, which provides better intrusion detection performance in UAV networks. The effectiveness of the technique is validated using NSL-KDD dataset. The DRL technique includes a reinforcement learning-based Deep Belief Network (DBN) developed for intrusion detection that allows the identification of intrusions into UAVs from networks. Furthermore, the BWO algorithm is used to determine the optimal values of the hyper-parameters within the proposed model. Analyses results show highly precise values which confirm the effectiveness of the approach.

In [118], a novel DRL-based approach for trajectory planning and interference rejection of an UAV in Internet of Things applications is presented. The radio environment is enhanced to suppress interfering signals and extend the desired signals using a re-configurable intelligent surface (RIS). Using Deep Deterministic Policy Gradient (DDPG) and Double Delay DDPG (TD3) models, the UAV autonomously learns its trajectory and RIS configuration based solely on changes in the received data rate. The simulation results show that the proposed DRL algorithms provide robust resistance to UAV interference, especially with the TD3 algorithm that provides faster and smoother convergence than the DDPG algorithm for larger RISs. Notably, this DRL-based approach offers important practical benefits by eliminating the need for explicit knowledge about RISs and disruptive channels.

Collision-free paths for multiple UAVs connected to cellular networks are explored by providing connectivity with Ground Base Stations (GBS) in the presence of a dynamic jammer in [119]. The problem is formulated as a sequential decision challenge in a discrete space, taking into account connectivity, collision avoidance, and kinematic constraints. Authors present an offline time difference (TD) learning algorithm that is combined with online SINR mapping as a solution approach. Specifically, an offline network is created and trained using the TD method to capture interactions between UAVs and the environment. Additionally, an online SINR mapping based DNN is designed and trained through supervised learning to capture the influence and changes caused by the block. The numerical results show that even without any mixer knowledge, the proposed algorithm achieves performance levels comparable to the ideal scenario with a perfect SINR map.

Secure communication between a UAV and a ground user in an urban environment is a significant requirement. was investigated. In [120], the stage includes several spies and a UAV jammer that generates artificial noise to disrupt spies' activities. The goal is to maximize stealth rates at the physical layer by co-opting the trajectory and transmission power of the UAVs. To handle the time-varying channel conditions, the problem is formulated as a Markov decision process. An advanced algorithm based on the double DQN is proposed to solve MDP. The simulation results show that the rapid convergence of the algorithm in various environments allows the UAV transmitter and UAV jammers to accurately determine the optimal locations to maximize information privacy rates. Furthermore, the performance comparison shows that the Double DQN (DDQN) algorithm outperforms the Q-learning and Deep Q-learning Network approaches. In another research, a secure evacuation system consisting of an end server, a ground control station, and a malicious eavesdropper UAV is presented [121]. The goal here is to maximize privacy by offering a UAV that may dynamically switch between scrambling and transfer modes. An algorithm based on a deep deterministic policy gradient (DDPG) is proposed to achieve the goal.

A physical layer security (PLS) in millimeter-wave rotary wing unmanned aerial vehicle communication using re-configurable intelligent surfaces is explained in [122]. Having multiple listeners and imperfect channel state information (CSI) may create a security problem in an UAV. To solve this problem, DRL is used to make real-time decisions in any time frame considering the dynamic UAV environment. To address continuous optimization variables, authors present a twin-twin-lag deep deterministic policy gradient (TTD3) approach, which maximizes the expected cumulative reward and improves the safe energy efficiency (SEE). The results demonstrate that the proposed method outperforms the traditional deep deterministic policy gradient twin DRL (TDDRL)-based approach.

A new framework with deep reinforcement learning for radio surveillance, in which a fixed-wing UAV is used to capture the radio fingerprint of a suspicious transmitter with the help of a benign Reconfigurable Intelligent Surface (RIS), is presented in [123]. The framework includes a newly designed TD3 model, which enables the UAV to learn both its trajectory and the RIS configuration that is based solely on the observed transmission rate of the suspicious transmitter. This research includes a fixed-wing UAV. Action and reward components are customized to suit the UAV's mobility constraint. Simulations demonstrate that the method offers the UAV reliable radio surveillance ability while it also keeps the desired distance from the UAV to the transmitter.

A covert communication system is used with intelligent reflective surfaces (IRS) in UAVs in [124]. The goal of the research is to improve the performance of confidential communications. Authors propose a new algorithm called Double Deep Q Network-based Orbit and Phase Optimization (TAP-DDQN). They can improve stealth communication performance by optimizing the 3D trajectory of the UAV and the phase configuration of the IRS. Through simulations, they demonstrate that TAP-DDQN outperforms reference solutions and leads to significant improvements in the stealth performance of the IRS-powered UAV stealth communication system.

In [134], authors present a framework that uses RL algorithms to detect intrusions in Mobile Unmanned Wireless Networks (MUWNs). The paper identifies three main categories of attacks on MUWNs, which are jamming, impersonation, and intrusion. Impersonation attacks involve attackers posing as UAVs within the MUWN to offer deceptive services or maliciously acquire data. Intrusions involve the direct upload of malicious software to the target MUWN. The authors present a case study which is based on DDPG algorithm to demonstrate how RL can be applied for intrusion detection.

Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm, known for its effectiveness in multi-agent situations, is proposed to overcome the obstacles presented by collaborative mixing and trajectory design in a scenario where multiple UAVs operate together [135]. UAVs fall into two categories. One category serves as aerial Base Stations (BSs) that transmit data to users, while the other one serves as jammer that emits fake sounds to disturb listeners on the ground. The proposed system assumes that UAVs have information about locations of both ground users and eavesdroppers. Training for both groups of UAVs is performed using MADDPG to dynamically change their position and maximize the combined safety ratio of all ground users. The safety ratio for a ground user is determined by subtracting the maximum acceptable signal-to-noise ratio for the entire ground eavesdropper from the received signal-to-noise ratio. Simulation results show that an efficient joint orbit design of UAVs is achieved by MADRL method. To improve learning efficiency and convergence, the Continuous Action Attention MADDPG (CAA-MADDPG) method is also presented. The simulation results show that CAA-MADDPG outperforms MADDPG and provides better reward performance.

In [136], a multi-UAV is used to receive signals from a ground station. There are some UAV jammers near the target that try to interfere with UAVs to reduce their received SINR. In this research, an UAV is considered as a unique system, such as an agent that is trained by RL algorithm, while jammers adjust their transmission powers and positions depending on predefined strategies. UAVs only use RSS and SINR to make decisions. A Q-learning-based model is trained by using RSS, SINR, and predicted trajectories of jammers, which are called knowledge-based RL. Although the location of the jammer is unknown, the agent can estimate the jammer trajectory based on the change of the RSS value with the distance and the flight inertia of the jammer.

A ground node sends confidential information to a legitimate UAV while a smart UAV eavesdropper is present that is able to adjust its position for an optimal eavesdropping [137]. The legitimate UAV and the eavesdropper are both treated as conflicting agents. The goal of the legitimate UAV is to maximize the total security while the eavesdropper's aim is to minimize security. The problem is redefined as a two-player zero-sum stochastic game (TZSG). Findings demonstrate that the legitimate UAV strategically chooses a communication link away from the Eavesdropper while the Eavesdropper seeks to intercept confidential information.

DRL helps UAVs to safely perform their duties by adapting themselves to the environment without the need for costly infrastructure changes. The DQN-based channel selection strategy may improve communication efficiency compared to the Q-learning-based strategy. In the literature, there are solutions against different attacks on UAVs using deep reinforcement methods, such as DQN, game theory, PHC, and actor–critical methods.

6. Analyses of security approaches

RL algorithms have been used to secure UAVs. However, the algorithms have been applied to ensure different security requirements in a various ways in UAV systems. In this research, we have compared the use of RL algorithms in UAV security according to different strategies. We consider strategies developed with RL algorithms for security UAVs. Table 2 contains a comparison of strategies, their limitations, and evaluation criteria. In general, it has been observed that RL algorithms provide better solutions to reduce security vulnerabilities ın UAVs. Specifically, RL algorithms are used for resource allocations, power sharing, trajectory optimizations, delay decisions, any other optimizations that can be used to detect attacks. Some limitations of RL algorithms are also discussed to show the applicability of the algorithms. Additionally, state–action-reward values are analyzed are related to the technical infrastructure of RL algorithms that are very significant for attacks. Table 3 contains RL algorithms that are used in UAV security in a comparative manner. Detailed explanations of RL algorithms that are used in UAVs security are given below.

- Q-Learning: This is a value-based algorithm that learns to predict the expected benefit from taking a certain action in a given situation. Q-learning algorithm uses a lookup table or Q table to store the expected rewards, which are known as Q values, for actions given a set of states. Q-learning requires a lot of memory when the number of cases and actions increase.
- DQN: Advantages of DQN include its ability to work with high-dimensional state spaces and to provide effective results for discrete action space. Limiting features of DQN are an overestimate of Q-values and inadequate policies. DQN also converges slowly. Moreover, it is affected by sample inefficiency in complex environments.
- Double DQN (DDQN): This algorithm uses two Q networks to reduce overestimation bias. Double DQN improves the learning efficiency and the stability of the algorithm [139]. DDQN algorithm provides more accurate and stable Q-value estimates than DQN. However, it requires additional computational resources due to maintaining two networks.
- TRPO: This is a policy-based algorithm that uses a trusted zone to improve the stability and the convergence. The trusted zone defines a constraint on how much the updated policy may deviate from the old policy. This circumstance avoids major policy changes that may cause instability [140]. Advantages of TRPO are managing continuous areas of actions and having access to consistent and reliable policy updates. However, it is computationally expensive due to the multiple iteration requirement.
- DDPG: It combines value-based and policy-based methods while maintaining separate networks of actors and critics. DDPG is an effective algorithm to handle continuous spaces. On the other hand, it requires careful tuning of hyper-parameters and scanning strategies [141]. Advantages of DDPG algorithm are a combination of Q-learning and policy gradient methods and it provide good performance in complex environments. It may also manage areas of continuous action. However, it suffers from instability during training stage. Therefore, it is difficult to train the algorithm in environments with an high-dimensional state space. Thus, DDPG requires a careful balance between exploration and exploitation.
- TD3: The core concept of Twin-Delayed Deep Deterministic Policy Gradient (TD3) is used to mitigate an overestimation bias that is inherent in Deep Q-Learning. Particularly, this case occurs in scenarios that contain discrete actions within an Actor-Critic domain [142]. TD3 improves the sample efficiency and stability. It is also robust in continuous motions. However, hyper-parameter tuning may be necessary for an optimal performance. Additionally, it may be sensitive to some environmental conditions. TTD3 is built on the success of TD3 and it may offer improved performance, but computing requirements are higher than TD3.
- MADDPG: This algorithm is designed for scenarios where multiple agents are learning simultaneously. MADDPG framework uses a training approach, where all agents are trained together, but each agent operates independently during the execution. In this context, each UAV acts as an individual agent. Observations and actions of all agents are used during the training stage. Moreover, each agent makes decisions based on its observations and the evaluated value during the execution stage [143]. MADDPG is suitable for multi-agent scenarios with a centralized training and a decentralized execution. It provides a stable training in complex and dynamic environments. Additionally, MADDPG may handle scenarios where agents have partial visibility. However, the computational complexity and training sensitivity to are main disadvantages of the algorithm.

Each algorithm has its own strengths and weaknesses. The choice of any algorithm depends on specific requirements of UAVs and characteristics of the environment. Hyper-parameter tuning, computational efficiency, adaptability to different scenarios, resource consumption, and high-dimensional state spaces are important facts that determine the choice of an algorithm in UAV systems.

7. Limitations and research challenges

7.1. Limitations

This section contains limitations of RL-based solutions available in the literature. Understanding the limitations will help researchers to create the most suitable and affordable countermeasures against various threats and attacks on UAVs. The limitations of RL-based solutions are shown in Fig. 13.

Limited training data. DRL algorithms require significant training data to learn effectively policies. However, it can be a challenging task to label training data for different attack scenarios and create countermeasures for security of UAVs. Moreover, it may not be feasible to obtain real-world data for all possible attack scenarios that will be used with DRL models.

Table 2
RL-based strategies for UAV security.

Ref.	RL techniques	Strategy	Limitations	Evaluation
[59]	DQN using CNN	Power Allocation	Highest computational complexity	Secrecy capacity, energy consumption cost
[115]	DQN using CNN	Power Control	Noise power, energy loss	Improves the signal-to-interference-plus-noise
[116]	Hotbooting PHC	Relay	Small region, high computation, communication overhead	Lower BER, higher utility
[117]	PHY-authentication, Q Learning	Spoofing detection	Dynamic radio environment	False alarm rate and miss detection rate
[118]	DDPG, TD3	Trajectory planning and jamming rejection	Non-convexity and sequential decision-making nature	Mission time, received data rate
[119]	TD Learning	Jamming-resilient path planning and trajectory designs	Connectivity and collisions in dynamic environment	Success Rate, Disconnection Rate, Collision Rate
[120]	DDQN	Optimizing UAV trajectory and transmission power	Large state space, multi-agent environment, computational complexity, complicated data transmission	Secrecy rates, convergence speed
[121]	DDPG	Relay decisions and offloading decisions	The generalization and the robustness	Secrecy sum-rate
[122]	TTD3	optimization of flight trajectory, UAV active beam-forming and RIS passive beam-forming	Multiple listeners and imperfect channel state information	Worst-case secrecy energy efficiency (SEE)
[123]	TTD3	RIS configuration and UAV navigation	Multi-UAV radio surveillance	Average eavesdropping success probability and average eavesdropping rate
[124]	TAP-DDQN	Improve the covert communication performance	More complex covert communication scenarios, multiple UAVs, dynamic eavesdroppers	Average covert rate
[125]	Hotbooting PHC	Relay	Transmission cost, the radio channel condition	Convergence speed, higher utility, reduce the exploration trials
[126]	PLASH PBE (Game Theory)	Decide the true position	Assumptions about rationality	Reduce the deviation between the estimated position and the true position
[128]	DRL-BWO	Intrusion detection	Scarcity of UAV-specific Data, data imbalance	High precision, recall, F-measure, and accuracy values
[138]	DQN using CNN	Relay Power	Service outages, energy exhaustion, secure relay	BER (Bit-Error-Rate), Save the UAV Energy Consumption
[135]	DDPG	Intrusion Detection	Computing-constrained mobile devices	Detection rate
[135]	MADDPG	Trajectory Design, Power Optimization	Large number of UAVs and jammers	Improve the learning efficiency and convergence, security rate
[136]	Q-Learning	Power strategy	Expensive and difficult to interact with real environment	Average utility, distance change of the mission UAV, reward value
[137]	TZSG	Optimizing the legitimate UAV trajectory, transmit power control and node scheduling	Uncontrollable mobility	Sum secrecy rate

Complexity of the UAV security environment. Securing UAV systems is a complex task because the environment may be dynamic and uncertain. RL algorithms generally assume a known static environment, which may not provide accurate security results for various UAV scenarios. Enemies can adapt and change their attack strategy. UAVs may be affected by external factors, such as the environment and weather conditions. Incorporating such complexities into RL models is a significant challenge that needs to be addressed.

Security issues. RL algorithms learn by trial and error, which may create security risks in UAV applications. Incorrect actions or policies learned from RL models may have disastrous consequences, including deadlocks or unauthorized access to critical systems. To counter security issues strict testing and verification procedures are required.

Adversarial attacks. RL models are vulnerable to adversarial attacks. Attackers can manipulate the training process or exploit vulnerabilities in RL algorithms. This circumstance may mislead RL models or simply

Table 3
Comparisons of RL approaches for UAV security.

Ref.	RL algorithms	States	Actions	Rewards	Attacks	Targets
[114]	DQN	PUs, SINR	Move or Stay	Utility	Jamming	Cognitive Radio Network
[115]	DQN	SINR	Power Transmission	SINR	Jamming	IoT Communication
[116]	PHC	Link Quality, SINR, BER	Relay Decision	Utility	Jamming	VANET
[117]	Q Learning	False Alarm, Miss Detection Rate of Authentication	Test Threshold	Utility	GPS Spoofing	Dynamic Radio Environments
[118]	DDPG, TD3	Relative Position, Velocity, SINR	Test Threshold	reflecting coefficients, acceleration	GPS Spoofing	Dynamic Radio Environments
[119]	TD Learning	Information vector, observable state of nearest UAVs, experienced SINR	Agent's current speed, orientation, the kinematic constraints	A function of the minimum distance to other UAVs	Jamming	Multi-UAV Cellular Network
[120]	MDP, Double DQN	Location of each UAV, legal UAV, and jammer UAV	moving speed, transmitting power	A function of the minimum distance to other UAVs	Jamming	Multi-UAV Cellular Network
[121]	DDPG	Coordinates of each UAV nodes, horizontal distance between legal UAV, and jammer UAV	Horizontal and vertical velocity of the legal UAV	Maximization of the secrecy sum-rate	Eavesdropping	Physical Layer Security
[122]	TTD3	Predicted comprehensive channel state information from UAV and eavesdroppers, local information,	Flying directions	Maximization of secrecy energy efficiency	Eavesdropping	Physical Layer Security
[123]	TD3	Relative position	Reflecting coefficients of the RIS and acceleration of the UAV	Eavesdropping rate maximization, destination, velocity constraint, altitude constraint	Eavesdropping	Radio Environment
[124]	TAP-DDQN	Coordinates of the current UAV	Horizontal and vertical coordinate, the binary indicator showing that the UAV is in service, continuous flight time	Maximization of power signal	Eavesdropping	Wireless Communication System
[134]	DDPG	UAV behaviors	Detection strategy at each time step (binary value)	System utility	Jamming, impersonation, intrusion attack	Mobile Unmanned Wireless Networks
[135]	MADDPG	Represent the current 3-D position of the agent, and the index of the objective GU	UAV's velocity projections on three orthogonal coordinate axes and the signal power of UAVs	Map limitation penalty, secure rate, power penalty and distance reward	Jamming	Mobile Unmanned Wireless Networks
[136]	Q-Learning	Maneuvering state (position and speed), channel state (allocate power for UAV an jammer)	Power distribution actions and maneuvering actions (stay, forward, backward, left, right, left front, right front, left back, and right back)	Minimum SINR requirement for the receiver, complete the mission, ability to flexibly	Jamming	Mobile Unmanned Wireless Networks
[137]	TZSG	Position and speed state and node scheduling state	Communication and the power control action, UAV action contain speed	Some constraints (max speed, min distance between two UAVs)	Eavesdropping	Physical Layer Security

cause undesired behaviors in executions of UAVs. Adversarial attacks targeting RL-based security solutions pose a significant challenge for UAV systems. Thus, robust defense mechanisms are needed to counter such attacks.

Transferability and generalization. RL models trained in a single environment may not be generalized to be used in new and unknown environments. UAV security scenarios may vary significantly. Training a RL model in one context may have poor performance results in a

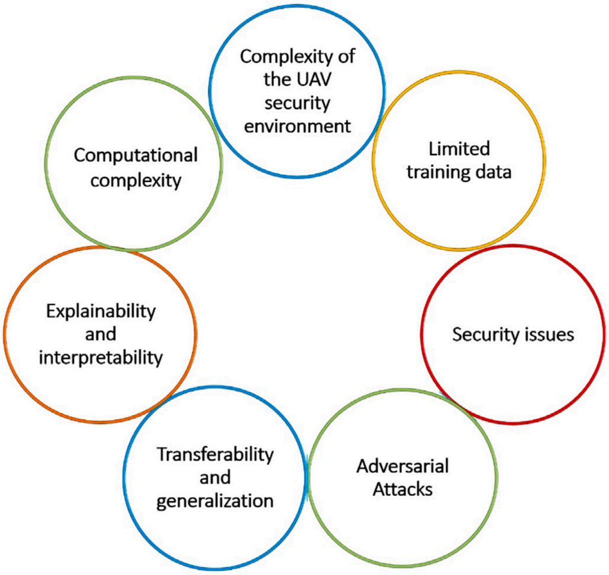

Complexity of the UAV security environment
Limited training data
Security issues
Adversarial Attacks
Transferability and generalization
Explainability and interpretability
Computational complexity

Fig. 13. Limitations of UAV security.

different context. The development of RL algorithms that may transfer information and adapt to new environments without extensive training is a challenge for UAV security.

Explainability and interpretability. The interpretation of DRL models is often hard process that makes it difficult to understand. A DRL based decision-making process of an agent is therefore hard to interpret. On the other hand, it is critical to have transparent and explainable models to identify potential vulnerabilities in UAVs to address them effectively. The development of interpretable RL models may provide insight about the logic of the UAV decision process, which may help to reduce vulnerabilities.

Computational complexity. DRL algorithms, especially those which are based on deep neural networks, may be computationally intensive and they require significant computational resources. If RL-based security solutions are applied on resource-constrained UAV platforms, the computational complexity may limit the solutions. Moreover, DRL based solutions may not provide real-time results, which may be unacceptable for UAVs that have critical tasks. Therefore, the development of lightweight and efficient RL algorithms suitable for UAV environments is a significant research challenge.

Limitations and security challenges must be addressed to counter security attacks on UAVs' systems. This situation requires collaborative efforts between researchers and industry experts. The efforts contain advanced data collection about UAV systems, new algorithmic developments, security assurances, correct interpretations and standard definitions, and guidelines for implementing RL-based security solutions in UAV applications.

7.2. Research challenges

This subsection contains analyses of potential research avenues about UAV vulnerabilities, threats, and attacks and their RL based solutions. In this research, the main focus of the research is improving the security of UAV communications and networks, where there are various threats and attacks. Fig. 14 shows the security challenges in UAV systems.

The use of deep reinforcement learning algorithms for UAV security is a relatively new and rapidly developing topic. Therefore, there are many avenues for future research challenges on the security of UAVs with RL. Some of the challenges are presented in [34]. For example, DRL requires users to report the local state at every time slot to determine the state in density networks. On the other hand, networks are expected to be deployed in a high density around a base station, where received signal strength indicators may be the same, which is a significant challenge for security of UAVs.

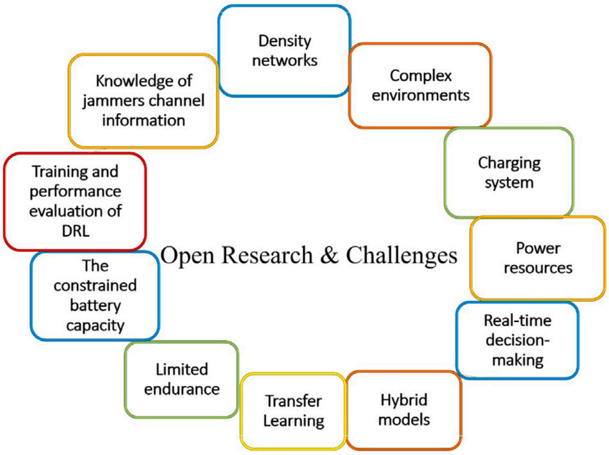

Density networks
Knowledge of jammers channel information
Complex environments
Training and performance evaluation of DRL
Charging system
Open Research & Challenges
Power resources
The constrained battery capacity
Real-time decision-making
Limited endurance
Transfer Learning
Hybrid models

Fig. 14. Security challenges for UAVs.

Knowledge about jammer's channel information is another research challenge. The reward function of an UAV requires a perfect knowledge about channel information of the jammer, which case is impossible in practice. However, the reward function need to be formulated. The lack of information about the channel makes it impractical to model the reward function.

Training and performance evaluations of DRL models are other challenging tasks for UAV security. The reason is that data are inaccessible in some wireless systems. Currently, many models use simulated datasets that are created with simulations. Creating UAV data with simulations may simplify the execution of an UAV system, which may not represent a real UAV system. Thus, the newly created RL based countermeasure may be effective against attacks.

A RL based security model created for complex environments is another research challenge. The complexity of an environment may prevent the RL model to adapt real-world scenarios. Therefore, designing DRL algorithms that support complex environments and provide robust real-world solutions is an important research challenge.

The lack of dynamic power load balancing within the UAV charging system introduces a significant performance overhead. This circumstance leads to problems including a power drain, a battery degradation, which results in the overall performance degradation [16]. Consequently, the need for optimized and efficient charging systems becomes a key challenge in this case. Addressing this challenge requires the development of solutions that may dynamically balance power loads, mitigate the associated problems, and optimize the overall performance of the charging system.

The limited battery capacity and the limited endurance of UAVs require improved power management strategies to ensure resilience in both survival and starvation scenarios [144]. The need to optimize power efficiency is a critical challenge. Additionally, the power management is critical to extend the overall lifetime of UAVs. Addressing this challenge involves improving the adaptation of power resources, thereby increasing endurance and making UAVs survivable under changing operational conditions.

Using transfer learning for an UAV security is another research challenge. Transfer learning involves pre-training a model and then adapting it to the new environment. In this model, training data are limited. Transfer learning may be used to improve the efficiency of DRL algorithms for the security of UAV systems. In this case, the challenge is to adequately pre-training the model.

Hybrid RL models have huge potential to counter threads and attacks in UAV systems. However, creating a hybrid RL based security solution for UAV systems is an important challenge. Combining DRL models with other machine learning algorithms are expected to increase the effectiveness and the efficiency of DRL based security solutions for UAV systems. On the other hand, combining many machine learning algorithms with DRL models increases the complexity of the security solution, where the probability of having vulnerabilities increases.

Real-time decision-making is another significant research challenge for the security of UAVs. In many cases, attackers use real-time interactions with UAV systems that require real-time responses from intrusion detection systems or simply from security countermeasures. Designing RL algorithms that make decisions in real-time according to some constraints of UAVs, such as hardware and the communication latency, is a significant challenge. Balancing the trade-off between the decision speed and the accuracy is critical for an effective security implementations.

The grand challenge is to secure the whole UAV systems with the most effective and the most appropriate security mechanisms. Conventional security mechanisms reduce the attack surface on UAV systems. However, they do not completely remove all threats or prevent attacks on UAV systems. DRL based security solutions help to reduce the attack surface on UAV systems. Similarly, DRL-based security solutions have research challenges that need to be solved before DRL based security solutions are applied on UAV systems.

8. Conclusion

This paper provides a systematic review about the security of unmanned aerial vehicles focusing on deep reinforcement learning based solutions. UAVs have been used more than ever in almost all environments. The vehicles are highly connected devices over various networks. Moreover, they run various algorithms to complete their tasks autonomously with limited resources, such as power and computing capabilities. This makes UAVs vulnerable to cyber attacks. Conventional security mechanisms do not provide an adequate security level for UAVs in many cases. Recently, machine learning based security solutions complement conventional security solutions.

In this article, a comprehensive review of UAV systems' communications is provided since many security threats and attacks are based on the network and the communication infrastructure of an UAV system. The general architecture of an UAV system is explained to make clear threat sources in UAVs. Moreover, types of communications in UAV systems with their security mechanisms are presented to show the current state of the art about communication options in UAVs. Additionally, security requirements of UAV systems are analyzed to show the need for reinforcement-based security solutions for UAVs.

Security attacks depend on security threats. A security attack is possible if there is at least one corresponding security threat. In this research, the main security threats and attacks on UAV systems are explained in a holistic way. Threats are first defined by considering environment properties and other facts to show the source of attacks in details. Then, common attacks on UAVs are explained with vulnerabilities they use. Furthermore, countermeasures in the literature against such attacks are examined to reveal requirements for defense mechanisms in UAV systems. Furthermore, traditional defense mechanisms applied on UAVs are investigated to extract deficiencies of the mechanisms.

Deep reinforcement learning based security solutions in the literature are investigated as complementary solutions to traditional security mechanisms applied on UAV systems. First, DRL are explained to show their capabilities and limitations. DRL based security solutions in literature implemented on UAVs are explained in detail to show the current state of the art about the solutions. Then, a comparative analysis of reinforcement learning based security solutions in UAV systems is presented to show the power of such solutions. Furthermore, we explore the limitations of DRL based security solutions. Finally, some interesting research challenges are discussed to improve the security of UAV systems.

CRediT authorship contribution statement

Burcu Sönmez Sarıkaya: Writing – original draft, Validation, Methodology, Conceptualization. **Şerif Bahtiyar:** Writing – review & editing, Validation, Supervision, Resources, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This work was supported by Research Fund of the Istanbul Technical University, Türkiye. Project Number: 45654

References

[1] Y. Mekdad, A. Aris, L. Babun, A.E. Fergougui, M. Conti, R. Lazzeretti, A.S. Uluagac, A survey on security and privacy issues of uavs, Comput. Netw. 224 (2023) 109626, http://dx.doi.org/10.1016/j.comnet.2023.109626.

[2] T. Patel, N. Salot, V. Parikh, A systematic literature review on security of unmanned aerial vehicle systems, 2022, arXiv:2212.05028, https://arxiv.org/abs/2212.05028.

[3] H.J. Hadi, Y. Cao, K.U. Nisa, A.M. Jamil, Q. Ni, A comprehensive survey on security, privacy issues and emerging defence technologies for uavs, J. Netw. Comput. Appl. 213 (2023) 103607, https://www.sciencedirect.com/science/article/pii/S1084804523000267.

[4] V. Hassija, V. Chamola, A. Agrawal, A. Goyal, N.C. Luong, D. Niyato, F.R. Yu, M. Guizani, Fast, reliable, and secure drone communication: A comprehensive survey, IEEE Commun. Surv. Tutor. 23 (4) (2021) 2802–2832, http://dx.doi.org/10.1109/comst.2021.3097916.

[5] J. Sharma, P.S. Mehra, Secure communication in iot-based uav networks: A systematic survey, Int. Things 23 (2023) 100883, https://api.semanticscholar.org/CorpusID:260101024.

[6] V. Hassija, V. Saxena, V. Chamola, A mobile data offloading framework based on a combination of blockchain and virtual voting, Softw. - Pract. Exp. 51 (12) (2021) 2428–2445, http://dx.doi.org/10.1002/spe.2786.

[7] S. Hafeez, A.R. Khan, M.M. Al-Quraan, L. Mohjazi, A. Zoha, M.A. Imran, Y. Sun, Blockchain-assisted uav communication systems: A comprehensive survey, IEEE Open J. Veh. Technol. 4 (2023) 558–580, http://dx.doi.org/10.1109/OJVT.2023.3295208.

[8] I. García-Magariño, R. Lacuesta, M. Rajarajan, J. Lloret, Security in networks of unmanned aerial vehicles for surveillance with an agent-based approach inspired by the principles of blockchain, Ad Hoc Netw. 86 (2019) 72–82, http://dx.doi.org/10.1016/j.adhoc.2018.11.010, https://www.sciencedirect.com/science/article/pii/S1570870518301689.

[9] M. Babiker Mohamed, O. Matthew Alofe, M. Ajmal Azad, H. Singh Lallie, K. Fatema, T. Sharif, A comprehensive survey on secure software-defined network for the internet of things, Trans. Emerg. Telecommun. Technol. 33 (1) (2022) http://dx.doi.org/10.1002/ett.4391.

[10] A. Gupta, S.K. Gupta, A study on secured unmanned aerial vehicle-based fog computing networks, SAE Int. J. Connect. Autom. Veh. 7 (12-07-02-0011) (2023) http://dx.doi.org/10.4271/12-07-02-0011.

[11] R.S. Sutton, A.G. Barto, et al., Introduction To Reinforcement Learning, vol. 135, MIT press Cambridge, 1998, https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf.

[12] A. Nandy, M. Biswas, Reinforcement Learning: With Open AI, TensorFlow and Keras using Python, A Press, 2017, https://link.springer.com/book/10.1007/978-1-4842-3285-9.

[13] Y. Bai, H. Zhao, X. Zhang, Z. Chang, R. Jäntti, K. Yang, Toward autonomous multi-uav wireless network: A survey of reinforcement learning-based approaches, Commun. Surv. Tuts. 25 (4) (2023) 3038–3067, http://dx.doi.org/10.1109/COMST.2023.3323344.

[14] C. Kwon, S. Yantek, I. Hwang, Real-time safety assessment of unmanned aircraft systems against stealthy cyber attacks, J. Aerosp. Inf. Syst. 13 (1) (2015) 27–45, http://dx.doi.org/10.2514/1.I010388.

[15] F. Ahmed, J. Mohanta, A. Keshari, P.S. Yadav, Recent advances in unmanned aerial vehicles: a review, Arab. J. Sci. Eng. 47 (7) (2022) 7963–7984, http://dx.doi.org/10.1007/s13369-022-06738-0.

[16] F. Tlili, L.C. Fourati, S. Ayed, B. Ouni, Investigation on vulnerabilities, threats and attacks prohibiting uavs charging and depleting uavs batteries: Assessments & countermeasures, Ad Hoc Netw. 129 (2022) 102805, http://dx.doi.org/10.1016/j.adhoc.2022.102805, https://www.sciencedirect.com/science/article/pii/S1570870522000208.

[17] R. Altawy, A.M. Youssef, Security, privacy, and safety aspects of civilian drones: A survey, ACM Trans. Cyber-Phys. Syst. 1 (2) (2016) http://dx.doi.org/10.1145/3001836.

[18] Y. Mekdad, A. Aris, L. Babun, A.E. Fergougui, M. Conti, R. Lazzeretti, A.S. Uluagac, A survey on security and privacy issues of uavs, 2021, arXiv:2109.14442, https://arxiv.org/abs/2109.14442.

[19] D. He, S. Chan, M. Guizani, Communication security of unmanned aerial vehicles, IEEE Wirel. Commun. 24 (4) (2017) 134–139, http://dx.doi.org/10.1109/MWC.2016.1600073WC.

[20] H. Hartenstein, L. Laberteaux, A tutorial survey on vehicular ad hoc networks, IEEE Commun. Mag. 46 (6) (2008) 164–171, http://dx.doi.org/10.1109/MCOM.2008.4539481.

[21] U. Challita, W. Saad, C. Bettstetter, Interference management for cellular-connected uavs: A deep reinforcement learning approach, IEEE Trans. Wireless Commun. 18 (4) (2019) 2125–2140, http://dx.doi.org/10.1109/twc.2019.2900035.

[22] Y. Zeng, R. Zhang, T.J. Lim, Wireless communications with unmanned aerial vehicles: opportunities and challenges, IEEE Commun. Mag. 54 (5) (2016) 36–42, http://dx.doi.org/10.1109/MCOM.2016.7470933.

[23] K.-Y. Tsao, T. Girdler, V.G. Vassilakis, A survey of cyber security threats and solutions for uav communications and flying ad-hoc networks, Ad Hoc Netw. 133 (2022) 102894, http://dx.doi.org/10.1016/j.adhoc.2022.102894, https://www.sciencedirect.com/science/article/pii/S1570870522000853.

[24] I. Jawhar, N. Mohamed, J. Al-Jaroodi, D.P. Agrawal, S. Zhang, Communication and networking of uav-based systems: Classification and associated architectures, J. Netw. Comput. Appl. 84 (2017) 93–108, http://dx.doi.org/10.1016/j.jnca.2017.02.008, https://www.sciencedirect.com/science/article/pii/S1084804517300814.

[25] L. Gupta, R. Jain, G. Vaszkun, Survey of important issues in uav communication networks, IEEE Commun. Surv. Tutor. 18 (2) (2016) 1123–1152, http://dx.doi.org/10.1109/COMST.2015.2495297.

[26] L. Krichen, M. Fourati, L.C. Fourati, Communication architecture for unmanned aerial vehicle system, in: N. Montavont, G.Z. Papadopoulos (Eds.), Ad-Hoc, Mobile, and Wireless Networks, Springer International Publishing, Cham, 2018, pp. 213–225, http://dx.doi.org/10.1007/978-3-030-00247-3_20.

[27] P. FIPS, 200, Minimum Security Requirements for Federal Information and Information Systems, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, 2006, http://dx.doi.org/10.6028/NIST.FIPS.200.

[28] A.Y. Javaid, W. Sun, V.K. Devabhaktuni, M. Alam, Cyber security threat analysis and modeling of an unmanned aerial vehicle system, in: 2012 IEEE Conference on Technologies for Homeland Security, HST, 2012, pp. 585–590, http://dx.doi.org/10.1109/THS.2012.6459914.

[29] V. Behzadan, Cyber-physical attacks on uas networks- challenges and open research problems, 2017, arXiv:1702.01251, https://arxiv.org/abs/1702.01251.

[30] T.M. Peake, Eavesdropping in Communication Networks, Cambridge University Press, 2005, pp. 13–37, http://dx.doi.org/10.1017/CBO9780511610363.004.

[31] D. He, S. Chan, M. Guizani, Drone-assisted public safety networks: The security aspect, IEEE Commun. Mag. 55 (8) (2017) 218–223, http://dx.doi.org/10.1109/MCOM.2017.1600799CM.

[32] Y. Mo, B. Sinopoli, Secure control against replay attacks, in: 2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2009, pp. 911–918, http://dx.doi.org/10.1109/ALLERTON.2009.5394956.

[33] A. Carrio, C. Sampedro, A. Rodriguez-Ramos, P.C. Cervera, A review of deep learning methods and applications for unmanned aerial vehicles, J. Sens. 2017 (2017) 3296874:1–3296874:13, https://api.semanticscholar.org/CorpusID:2958701.

[34] N.C. Luong, D.T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, D.I. Kim, Applications of deep reinforcement learning in communications and networking: A survey, IEEE Commun. Surv. Tutor. 21 (4) (2019) 3133–3174, http://dx.doi.org/10.1109/COMST.2019.2916583.

[35] J.-A. Maxa, M.S.B. Mahmoud, N. Larrieu, Survey on uaanet routing protocols and network security challenges, Ad Hoc Sens. Wirel. Netw. 37 (2017) 231–320, https://api.semanticscholar.org/CorpusID:11587162.

[36] A. Chriki, H. Touati, H. Snoussi, F. Kamoun, Fanet: Communication, mobility models and security issues, Comput. Netw. 163 (2019) 106877, http://dx.doi.org/10.1016/j.comnet.2019.106877, https://www.sciencedirect.com/science/article/pii/S1389128618309034.

[37] C.-L. Chen, Y.-Y. Deng, W. Weng, C.-H. Chen, Y.-J. Chiu, C.-M. Wu, A traceable and privacy-preserving authentication for uav communication control system, Electronics 9 (1) (2020) http://dx.doi.org/10.3390/electronics9010062, https://www.mdpi.com/2079-9292/9/1/62.

[38] I. Cervesato, The dolev-yao intruder is the most powerful attacker, in: 16th Annual Symposium on Logic in Computer Science—LICS, Vol. 1, Citeseer, 2001.

[39] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, T. Melodia, Machine learning for wireless communications in the internet of things: A comprehensive survey, Ad Hoc Netw. 93 (2019) 101913, http://dx.doi.org/10.1016/j.adhoc.2019.101913, https://www.sciencedirect.com/science/article/pii/S1570870519300812.

[40] P. Yi, Z. Dai, S. Zhang, Y. Zhong, et al., A new routing attack in mobile ad hoc networks, Int. J. Inf. Technol. 11 (2) (2005) 83–94.

[41] T. Kavitha, D. Sridharan, Security vulnerabilities in wireless sensor networks: A survey, J. Inf. Assur. Secur. 5 (1) (2010) 31–44.

[42] A. Rovira-Sugranes, A. Razi, F. Afghah, J. Chakareski, A review of ai-enabled routing protocols for uav networks: Trends, challenges, and future outlook, Ad Hoc Netw. 130 (2022) 102790, http://dx.doi.org/10.1016/j.adhoc.2022.102790, https://www.sciencedirect.com/science/article/pii/S1570870522000087.

[43] H. Nawaz, H.M. Ali, A.A. Laghari, Uav communication networks issues: a review, Arch. Comput. Methods Eng. (2020) 1–21, http://dx.doi.org/10.1007/s11831-020-09418-0.

[44] D. Westhoff, B. Lamparter, C. Paar, A. Weimerskirch, On digital signatures in ad hoc networks, Eur. Trans. Telecommun. 16 (5) (2005) 411–425, http://dx.doi.org/10.1002/ett.1061.

[45] Saifullah, Z. Ren, K. Hussain, M. Faheem, K-means online-learning routing protocol (k-morp) for unmanned aerial vehicles (uav) adhoc networks, Ad Hoc Netw. 154 (2024) 103354, http://dx.doi.org/10.1016/j.adhoc.2023.103354, https://www.sciencedirect.com/science/article/pii/S1570870523002743.

[46] A. Zear, V. Ranga, Uavs assisted network partition detection and connectivity restoration in wireless sensor and actor networks, Ad Hoc Netw. 130 (2022) 102823, http://dx.doi.org/10.1016/j.adhoc.2022.102823, https://www.sciencedirect.com/science/article/pii/S1570870522000336.

[47] A. Rovira-Sugranes, A. Razi, F. Afghah, J. Chakareski, A review of ai-enabled routing protocols for uav networks: Trends, challenges, and future outlook, Ad Hoc Netw. 130 (2022) 102790, http://dx.doi.org/10.1016/j.adhoc.2022.102790, https://www.sciencedirect.com/science/article/pii/S1570870522000087.

[48] J. Xue, Q. Wu, H. Zhang, Cost optimization of uav-mec network calculation offloading: A multi-agent reinforcement learning method, Ad Hoc Netw. 136 (2022) 102981, http://dx.doi.org/10.1016/j.adhoc.2022.102981, https://www.sciencedirect.com/science/article/pii/S1570870522001548.

[49] M.M. Kermani, M. Zhang, A. Raghunathan, N.K. Jha, Emerging frontiers in embedded security, in: 2013 26th International Conference on VLSI Design and 2013 12th International Conference on Embedded Systems, 2013, pp. 203–208, http://dx.doi.org/10.1109/VLSID.2013.222.

[50] A.M. Nia, M. Mozaffari-Kermani, S. Sur-Kolay, A. Raghunathan, N.K. Jha, Energy-efficient long-term continuous personal health monitoring, IEEE Trans. Multi-Scale Comput. Syst. 1 (2) (2015) 85–98, http://dx.doi.org/10.1109/TMSCS.2015.2494021.

[51] F.A. Yerlikaya, Şerif Bahtiyar, Data poisoning attacks against machine learning algorithms, Expert Syst. Appl. 208 (2022) 118101, http://dx.doi.org/10.1016/j.eswa.2022.118101, https://www.sciencedirect.com/science/article/pii/S0957417422012933.

[52] M. Mozaffari-Kermani, S. Sur-Kolay, A. Raghunathan, N.K. Jha, Systematic poisoning attacks on and defenses for machine learning in healthcare, IEEE J. Biomed. Health Inf. 19 (6) (2015) 1893–1905, http://dx.doi.org/10.1109/JBHI.2014.2344095.

[53] C.G.L. Krishna, R.R. Murphy, A review on cybersecurity vulnerabilities for unmanned aerial vehicles, in: 2017 IEEE International Symposium on Safety, Security and Rescue Robotics, SSRR, 2017, pp. 194–199, http://dx.doi.org/10.1109/SSRR.2017.8088163.

[54] K. Mansfield, T. Eveleigh, T.H. Holzer, S. Sarkani, Unmanned aerial vehicle smart device ground control station cyber security threat model, in: 2013 IEEE International Conference on Technologies for Homeland Security, HST, 2013, pp. 722–728, http://dx.doi.org/10.1109/THS.2013.6699093.

[55] P. Boccadoro, D. Striccoli, L.A. Grieco, An extensive survey on the internet of drones, Ad Hoc Netw. 122 (2021) 102600, http://dx.doi.org/10.1016/j.adhoc.2021.102600, https://www.sciencedirect.com/science/article/pii/S1570870521001335.

[56] N.M. Rodday, R. d.O. Schmidt, A. Pras, Exploring security vulnerabilities of unmanned aerial vehicles, in: NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium, 2016, pp. 993–994, http://dx.doi.org/10.1109/NOMS.2016.7502939.

[57] J. Su, J. He, P. Cheng, J. Chen, A stealthy gps spoofing strategy for manipulating the trajectory of an unmanned aerial vehicle, IFAC-PapersOnLine 49 (22) (2016) 291–296, http://dx.doi.org/10.1016/j.ifacol.2016.10.412, 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2016. https://www.sciencedirect.com/science/article/pii/S2405896316319991.

[58] A.J. Kerns, D.P. Shepard, J.A. Bhatti, T.E. Humphreys, Unmanned aircraft capture and control via gps spoofing, J. Field Robot. 31 (4) (2014) 617–636, http://dx.doi.org/10.1002/rob.21513.

[59] L. Xiao, C. Xie, M. Min, W. Zhuang, User-centric view of unmanned aerial vehicle transmission against smart attacks, IEEE Trans. Veh. Technol. 67 (2018) 3420–3430, https://api.semanticscholar.org/CorpusID:4941779.

[60] A. Shafique, A. Mehmood, M. Elhadef, Survey of security protocols and vulnerabilities in unmanned aerial vehicles, IEEE Access 9 (2021) 46927–46948, http://dx.doi.org/10.1109/ACCESS.2021.3066778.

[61] A. Abbaspour, K.K. Yen, S. Noei, A. Sargolzaei, Detection of fault data injection attack on uav using adaptive neural network, Procedia Comput. Sci. 95 (2016) 193–200, http://dx.doi.org/10.1016/j.procs.2016.09.312, complex Adaptive Systems Los Angeles, CA November (2016) 2-4. https://www.sciencedirect.com/science/article/pii/S1877050916324851.

[62] I. Samy, I. Postlethwaite, D.-W. Gu, Survey and application of sensor fault detection and isolation schemes, Control Eng. Pract. 19 (7) (2011) 658–674, http://dx.doi.org/10.1016/j.conengprac.2011.03.002, https://www.sciencedirect.com/science/article/pii/S0967066111000414.

[63] H. Talebi, R. Patel, An intelligent fault detection and recovery scheme for reaction wheel actuator of satellite attitude control systems, in: 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006, pp. 3282–3287, http://dx.doi.org/10.1109/CACSD-CCA-ISIC.2006.4777164.

[64] Q. Shen, B. Jiang, P. Shi, C.-C. Lim, Novel neural networks-based fault tolerant control scheme with fault alarm, IEEE Trans. Cybern. 44 (11) (2014) 2190–2201, http://dx.doi.org/10.1109/TCYB.2014.2303131.

[65] V. Desnitsky, N. Rudavin, I. Kotenko, Modeling and evaluation of battery depletion attacks on unmanned aerial vehicles in crisis management systems, in: I. Kotenko, C. Badica, V. Desnitsky, D. El Baz, M. Ivanovic (Eds.), Intelligent Distributed Computing XIII, Springer International Publishing, Cham, 2020, pp. 323–332, http://dx.doi.org/10.1007/978-3-030-32258-8_38.

[66] S. Yi, P. Naldurg, R. Kravets, Security-aware ad hoc routing for wireless networks, in: Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing, MobiHoc '01, Association for Computing Machinery, New York, NY, USA, 2001, pp. 299–302, http://dx.doi.org/10.1145/501449.501464.

[67] C. Harsch, A. Festag, P. Papadimitratos, Secure position-based routing for vanets, in: 2007 IEEE 66th Vehicular Technology Conference, 2007, pp. 26–30, http://dx.doi.org/10.1109/VETECF.2007.22.

[68] S. Desilva, R. Boppana, Mitigating malicious control packet floods in ad hoc networks, in: IEEE Wireless Communications and Networking Conference, Vol. 4, 2005, pp. 2112–2117, http://dx.doi.org/10.1109/WCNC.2005.1424844.

[69] X. Sun, D.W.K. Ng, Z. Ding, Y. Xu, Z. Zhong, Physical layer security in uav systems: Challenges and opportunities, IEEE Wirel. Commun. 26 (5) (2019) 40–47, http://dx.doi.org/10.1109/MWC.001.1900028.

[70] S.K. Das, K. Kant, N. Zhang (Eds.), Handbook on Securing Cyber-Physical Critical Infrastructure, Morgan Kaufmann, Boston, 2012, pp. xi–xv, http://dx.doi.org/10.1016/B978-0-12-415815-3.00034-0, [Contributors], https://www.sciencedirect.com/science/article/pii/B9780124158153000340.

[71] Y. Li, I. Frasure, A.A. Ikusan, J. Zhang, R. Dai, Vulnerability assessment for unmanned systems autonomy services architecture, in: M.H. Au, S.M. Yiu, J. Li, X. Luo, C. Wang, A. Castiglione, K. Kluczniak (Eds.), Network and System Security, Springer International Publishing, Cham, 2018, pp. 266–276, http://dx.doi.org/10.1007/978-3-030-02744-5_20.

[72] Y. Li, R. Dai, J. Zhang, Morphing communications of cyber–physical systems towards moving-target defense, in: 2014 IEEE International Conference on Communications, ICC, 2014, pp. 592–598, http://dx.doi.org/10.1109/ICC.2014.6883383.

[73] M. Anastasova, M. Bisheh-Niasar, H. Seo, R. Azarderakhsh, M.M. Kermani, Efficient and side-channel resistant design of high-security ed448 on arm cortex-m4, in: 2022 IEEE International Symposium on Hardware Oriented Security and Trust, HOST, 2022, pp. 93–96, http://dx.doi.org/10.1109/HOST54066.2022.9839742.

[74] M. Bisheh-Niasar, R. Azarderakhsh, M. Mozaffari-Kermani, Cryptographic accelerators for digital signature based on ed25519, IEEE Trans. Very Large Scale Integr. (VLSI) Syst. 29 (7) (2021) 1297–1305, http://dx.doi.org/10.1109/TVLSI.2021.3077885.

[75] M. Anastasova, R.E. Khatib, A. Laclaustra, R. Azarderakhsh, M.M. Kermani, Highly optimized curve448 and ed448 design in wolfssl and side-channel evaluation on cortex-m4, in: 2023 IEEE Conference on Dependable and Secure Computing, DSC, 2023, pp. 1–8, http://dx.doi.org/10.1109/DSC61021.2023.10354154.

[76] V.L. Thing, J. Wu, Autonomous vehicle security: A taxonomy of attacks and defences, in: 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2016, pp. 164–170, http://dx.doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.52.

[77] M. Zhu, S. Martínez, On the performance analysis of resilient networked control systems under replay attacks, IEEE Trans. Autom. Control 59 (3) (2014) 804–808, http://dx.doi.org/10.1109/TAC.2013.2279896.

[78] B. Chen, D.W.C. Ho, G. Hu, L. Yu, Secure fusion estimation for bandwidth constrained cyber–physical systems under replay attacks, IEEE Trans. Cybern. 48 (6) (2018) 1862–1876, http://dx.doi.org/10.1109/TCYB.2017.2716115.

[79] F. Miao, M. Pajic, G.J. Pappas, Stochastic game approach for replay attack detection, in: 52nd IEEE Conference on Decision and Control, 2013, pp. 1854–1859, http://dx.doi.org/10.1109/CDC.2013.6760152.

[80] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, G. Pantziou, A survey on jamming attacks and countermeasures in wsns, IEEE Commun. Surv. Tutor. 11 (4) (2009) 42–56, http://dx.doi.org/10.1109/SURV.2009.090404.

[81] V. Chamola, P. Kotesh, A. Agarwal, Naren, N. Gupta, M. Guizani, A comprehensive review of unmanned aerial vehicle attacks and neutralization techniques, Ad Hoc Netw. 111 (2021) 102324, http://dx.doi.org/10.1016/j.adhoc.2020.102324, https://www.sciencedirect.com/science/article/pii/S1570870520306788.

[82] Y. Guan, X. Ge, Distributed attack detection and secure estimation of networked cyber–physical systems against false data injection attacks and jamming attacks, IEEE Trans. Signal Inf. Process. Netw. 4 (1) (2018) 48–59, http://dx.doi.org/10.1109/TSIPN.2017.2749959.

[83] O. Osanaiye, A.S. Alfa, G.P. Hancke, A statistical approach to detect jamming attacks in wireless sensor networks, Sensors 18 (6) (2018) http://dx.doi.org/10.3390/s18061691, https://www.mdpi.com/1424-8220/18/6/1691.

[84] T.E. Humphreys, Statement on the vulnerability of civil unmanned aerial vehicles and other systems to civil gps, 2012, https://api.semanticscholar.org/CorpusID:35123487.

[85] L. He, W. Li, C. Guo, R. Niu, Civilian unmanned aerial vehicle vulnerability to gps spoofing attacks, in: 2014 Seventh International Symposium on Computational Intelligence and Design, Vol. 2, 2014, pp. 212–215, http://dx.doi.org/10.1109/ISCID.2014.131.

[86] H. Hu, N. Wei, A study of gps jamming and anti-jamming, in: 2009 2nd International Conference on Power Electronics and Intelligent Transportation System, PEITS, Vol. 1, 2009, pp. 388–391, http://dx.doi.org/10.1109/PEITS.2009.5406988.

[87] S.-H. Seo, B.-H. Lee, S.-H. Im, G.-I. Jee, Effect of spoofing on unmanned aerial vehicle using counterfeited gps signal, J. Position. Navig. Timing 4 (2) (2015) 57–65, http://dx.doi.org/10.11003/JPNT.2015.4.2.057.

[88] M.L. Psiaki, B.W. O'Hanlon, J.A. Bhatti, D.P. Shepard, T.E. Humphreys, Gps spoofing detection via dual-receiver correlation of military signals, IEEE Trans. Aerosp. Electron. Syst. 49 (4) (2013) 2250–2267, http://dx.doi.org/10.1109/TAES.2013.6621814.

[89] Q. Zou, S. Huang, F. Lin, M. Cong, Detection of gps spoofing based on uav model estimation, in: IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, 2016, pp. 6097–6102, http://dx.doi.org/10.1109/IECON.2016.7793069.

[90] C. Pu, Y. Li, Lightweight authentication protocol for unmanned aerial vehicles using physical unclonable function and chaotic system, in: 2020 IEEE International Symposium on Local and Metropolitan Area Networks, LANMAN, 2020, pp. 1–6, http://dx.doi.org/10.1109/LANMAN49260.2020.9153239.

[91] M. Bellare, C. Namprempre, Authenticated encryption: Relations among notions and analysis of the generic composition paradigm, in: Advances in Cryptology—ASIACRYPT 2000: 6th International Conference on the Theory and Application of Cryptology and Information Security Kyoto, Japan, December 3–7, 2000 Proceedings 6, Springer, 2000, pp. 531–545, http://dx.doi.org/10.1007/s00145-008-9026-x.

[92] S. He, Q. Wu, J. Liu, W. Hu, B. Qin, Y.-N. Li, Secure communications in unmanned aerial vehicle network, in: Information Security Practice and Experience: 13th International Conference, ISPEC 2017, Melbourne, VIC, Australia, December 13–15, 2017, Proceedings 13, Springer, 2017, pp. 601–620, http://dx.doi.org/10.1007/978-3-319-72359-4_37.

[93] D.S.C. Putranto, A.K. Aji, B. Wahyudono, Design and implementation of secure transmission on internet of drones, in: 2019 IEEE 6th Asian Conference on Defence Technology, ACDT, 2019, pp. 128–135, http://dx.doi.org/10.1109/ACDT47198.2019.9072714.

[94] K.D. Wesson, T.E. Humphreys, B.L. Evans, Can cryptography secure next generation air traffic surveillance?, 2014, https://api.semanticscholar.org/CorpusID:21207906.

[95] Pan Wei-jun, Feng Ziliang, Wang Yang, Ads-b data authentication based on ecc and x.509 certificate, 2012, https://api.semanticscholar.org/CorpusID:16106662.

[96] L. Teng, M. Jianfeng, F. Pengbin, M. Yue, M. Xindi, Z. Jiawei, C. Gao, L. Di, Lightweight security authentication mechanism towards uav networks, in: 2019 International Conference on Networking and Network Applications (NaNA), 2019, pp. 379–384, http://dx.doi.org/10.1109/NaNA.2019.00072.

[97] J. Srinivas, A.K. Das, N. Kumar, J.J.P.C. Rodrigues, Tcalas: Temporal credential-based anonymous lightweight authentication scheme for internet of drones environment, IEEE Trans. Veh. Technol. 68 (7) (2019) 6903–6916, http://dx.doi.org/10.1109/TVT.2019.2911672.

[98] Z. Ali, S.A. Chaudhry, M.S. Ramzan, F. Al-Turjman, Securing smart city surveillance: A lightweight authentication mechanism for unmanned vehicles, IEEE Access 8 (2020) 43711–43724, http://dx.doi.org/10.1109/ACCESS.2020.2977817.

[99] B. Koziel, R. Azarderakhsh, M. Mozaffari Kermani, D. Jao, Post-quantum cryptography on fpga based on isogenies on elliptic curves, IEEE Trans. Circuits Syst. I. Regul. Pap. 64 (1) (2017) 86–99, http://dx.doi.org/10.1109/TCSI.2016.2611561.

[100] M. Bisheh-Niasar, R. Azarderakhsh, M. Mozaffari-Kermani, High-speed ntt-based polynomial multiplication accelerator for post-quantum cryptography, in: 2021 IEEE 28th Symposium on Computer Arithmetic, ARITH, 2021, pp. 94–101, http://dx.doi.org/10.1109/ARITH51176.2021.00028.

[101] M. Anastasova, R. Azarderakhsh, M.M. Kermani, Fast strategies for the implementation of sike round 3 on arm cortex-m4, IEEE Trans. Circuits Syst. I. Regul. Pap. 68 (10) (2021) 4129–4141, http://dx.doi.org/10.1109/TCSI.2021.3096916.

[102] Z. Yang, H. Alfauri, B. Farkiani, R. Jain, R.D. Pietro, A. Erbad, A survey and comparison of post-quantum and quantum blockchains, IEEE Commun. Surv. Tutor. 26 (2) (2024) 967–1002, http://dx.doi.org/10.1109/COMST.2023.3325761.

[103] G. Choudhary, V. Sharma, I. You, K. Yim, I.-R. Chen, J.-H. Cho, Intrusion detection systems for networked unmanned aerial vehicles: A survey, in: 2018 14th International Wireless Communications & Mobile Computing Conference, IWCMC, 2018, pp. 560–565, http://dx.doi.org/10.1109/IWCMC.2018.8450305.

[104] J. Xu, Z. Deng, Q. Song, Q. Chi, T. Wu, Y. Huang, D. Liu, M. Gao, Multi-uav counter-game model based on uncertain information, Appl. Math. Comput. 366 (2020) 124684, http://dx.doi.org/10.1016/j.amc.2019.124684, https://www.sciencedirect.com/science/article/pii/S0096300319306769.

[105] J.-P. Condomines, R. Zhang, N. Larrieu, Network intrusion detection system for uav ad-hoc communication: From methodology design to real test validation, Ad Hoc Netw. 90 (2019) 101759, http://dx.doi.org/10.1016/j.adhoc.2018.09.004, recent advances on security and privacy in Intelligent Transportation Systems. https://www.sciencedirect.com/science/article/pii/S1570870518306541.

[106] J. Whelan, T. Sangarapillai, O. Minawi, A. Almehmadi, K. El-Khatib, Novelty-based intrusion detection of sensor attacks on unmanned aerial vehicles, in: Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks, Q2SWinet '20, Association for Computing Machinery, New York, NY, USA, 2020, pp. 23–28, http://dx.doi.org/10.1145/3416013.3426446.

[107] T.M. Hoang, N.M. Nguyen, T.Q. Duong, Detection of eavesdropping attack in uav-aided wireless systems: Unsupervised learning with one-class svm and k-means clustering, IEEE Wirel. Commun. Lett. 9 (2) (2020) 139–142, http://dx.doi.org/10.1109/LWC.2019.2945022.

[108] M.P. Arthur, Detecting signal spoofing and jamming attacks in uav networks using a lightweight ids, in: 2019 International Conference on Computer, Information and Telecommunication Systems, CITS, 2019, pp. 1–5, http://dx.doi.org/10.1109/CITS.2019.8862148.

[109] M.B. Bejiga, A. Zeggada, A. Nouffidj, F. Melgani, A convolutional neural network approach for assisting avalanche search and rescue operations with uav imagery, Remote Sens. 9 (2) (2017) http://dx.doi.org/10.3390/rs9020100, https://www.mdpi.com/2072-4292/9/2/100.

[110] B. Yang, E.T. Matson, A.H. Smith, J.E. Dietz, J.C. Gallagher, Uav detection system with multiple acoustic nodes using machine learning models, in: 2019 Third IEEE International Conference on Robotic Computing, IRC, 2019, pp. 493–498, http://dx.doi.org/10.1109/IRC.2019.00103.

[111] A. Yazdinejad, R.M. Parizi, A. Dehghantanha, H. Karimipour, Federated learning for drone authentication, Ad Hoc Netw. 120 (2021) 102574, http://dx.doi.org/10.1016/j.adhoc.2021.102574, https://www.sciencedirect.com/science/article/pii/S1570870521001165.

[112] C. Pu, P. Zhu, Defending against flooding attacks in the internet of drones environment, in: 2021 IEEE Global Communications Conference, GLOBECOM, 2021, pp. 1–6, http://dx.doi.org/10.1109/GLOBECOM46510.2021.9686017.

[113] J.-P. Condomines, R. Zhang, N. Larrieu, Network intrusion detection system for uav ad-hoc communication: From methodology design to real test validation, Ad Hoc Netw. 90 (2019) 101759, http://dx.doi.org/10.1016/j.adhoc.2018.09.004, recent advances on security and privacy in Intelligent Transportation Systems. https://www.sciencedirect.com/science/article/pii/S1570870518306541.

[114] G. Han, L. Xiao, H.V. Poor, Two-dimensional anti-jamming communication based on deep reinforcement learning, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2017, pp. 2087–2091, http://dx.doi.org/10.1109/ICASSP.2017.7952524.

[115] Y. Chen, Y. Li, D. Xu, L. Xiao, Dqn-based power control for iot transmission against jamming, in: 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), 2018, pp. 1–5, http://dx.doi.org/10.1109/VTCSpring.2018.8417695.

[116] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, W. Zhuang, Uav relay in vanets against smart jamming with reinforcement learning, IEEE Trans. Veh. Technol. 67 (5) (2018) 4087–4097, http://dx.doi.org/10.1109/TVT.2018.2789466.

[117] L. Xiao, Y. Li, G. Liu, Q. Li, W. Zhuang, Spoofing detection with reinforcement learning in wireless networks, in: 2015 IEEE Global Communications Conference, GLOBECOM, 2015, pp. 1–5, http://dx.doi.org/10.1109/GLOCOM.2015.7417078.

[118] S. Hu, X. Yuan, W. Ni, X. Wang, A. Jamalipour, Ris-assisted jamming rejection and path planning for uav-borne iot platform: A new deep reinforcement learning framework, 2023, arXiv:2302.04994. https://arxiv.org/abs/2302.04994.

[119] X. Wang, M. Cenk Gursoy, T. Erpek, Y.E. Sagduyu, Jamming-resilient path planning for multiple uavs via deep reinforcement learning, in: 2021 IEEE International Conference on Communications Workshops (ICC Workshops), 2021, pp. 1–6, http://dx.doi.org/10.1109/ICCWorkshops50388.2021.9473587.

[120] Z. Qian, Z. Deng, C. Cai, H. Li, Reinforcement learning based dual-uav trajectory optimization for secure communication, Electronics 12 (9) (2023) http://dx.doi.org/10.3390/electronics12092008, https://www.mdpi.com/2079-9292/12/9/2008.

[121] S. Yoo, S. Jeong, J. Kang, Hybrid uav-enabled secure offloading via deep reinforcement learning, IEEE Wirel. Commun. Lett. 12 (6) (2023) 972–976, http://dx.doi.org/10.1109/LWC.2023.3254554.

[122] M.-L. Tham, Y.J. Wong, A. Iqbal, N.B. Ramli, Y. Zhu, T. Dagiuklas, Deep reinforcement learning for secrecy energy- efficient uav communication with reconfigurable intelligent surface, in: 2023 IEEE Wireless Communications and Networking Conference, WCNC, 2023, pp. 1–6, http://dx.doi.org/10.1109/WCNC55385.2023.10118891.

[123] X. Yuan, S. Hu, W. Ni, X. Wang, A. Jamalipour, Deep reinforcement learning-driven reconfigurable intelligent surface-assisted radio surveillance with a fixed-wing uav, IEEE Trans. Inf. Forensics Secur. 18 (2023) 4546–4560, http://dx.doi.org/10.1109/TIFS.2023.3297021.

[124] S. Bi, L. Hu, Q. Liu, J. Wu, R. Yang, L. Wu, Deep reinforcement learning for irs-assisted uav covert communications, China Commun. 20 (12) (2023) 131–141, http://dx.doi.org/10.23919/JCC.ea.2022-0336.202302.

[125] X. Lu, D. Xu, L. Xiao, L. Wang, W. Zhuang, Anti-jamming communication game for uav-aided vanets, in: GLOBECOM 2017-2017 IEEE Global Communications Conference, 2017, pp. 1–6, http://dx.doi.org/10.1109/GLOCOM.2017.8253987.

[126] T. Zhang, Q. Zhu, Strategic defense against deceptive civilian gps spoofing of unmanned aerial vehicles, in: S. Rass, B. An, C. Kiekintveld, F. Fang, S. Schauer (Eds.), Decision and Game Theory for Security, Springer International Publishing, Cham, 2017, pp. 213–233, http://dx.doi.org/10.1007/978-3-319-68711-7_12.

[127] H. Sedjelmaci, A. Boudguiga, I.B. Jemaa, S.M. Senouci, An efficient cyber defense framework for uav-edge computing network, Ad Hoc Netw. 94 (2019) 101970, http://dx.doi.org/10.1016/j.adhoc.2019.101970, https://www.sciencedirect.com/science/article/pii/S1570870519302136.

[128] V. Praveena, A. Vijayaraj, P. Chinnasamy, I. Ali, R. Alroobaea, S.Y. Alyahyan, M.A. Raza, Optimal deep reinforcement learning for intrusion detection in uavs, Comput. Mater. Continua 70 (2) (2022) 2639–2653, http://dx.doi.org/10.32604/cmc.2022.020066, http://www.techscience.com/cmc/v70n2/44676.

[129] A.R. Svaigen, A. Boukerche, L.B. Ruiz, A.A. Loureiro, Trajectory matters: Impact of jamming attacks over the drone path planning on the internet of drones, Ad Hoc Netw. 146 (2023) 103179, http://dx.doi.org/10.1016/j.adhoc.2023.103179, https://www.sciencedirect.com/science/article/pii/S1570870523000999.

[130] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: a survey, J. Artif. Int. Res. 4 (1) (1996) 237–285.

[131] L.P. Kaelbling, M.L. Littman, A.W. Moore, An introduction to reinforcement learning, in: The Biology and Technology of Intelligent Autonomous Agents, Springer, 1995, pp. 90–127.

[132] H. Kimura, M. Yamamura, S. Kobayashi, Reinforcement learning by stochastic hill climbing on discounted reward, in: A. Prieditis, S. Russell (Eds.), Machine Learning Proceedings 1995, Morgan Kaufmann, San Francisco (CA), 1995, pp. 295–303, http://dx.doi.org/10.1016/B978-1-55860-377-6.50044-X, https://www.sciencedirect.com/science/article/pii/B978155860377650044X.

[133] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (10) (2010) 1345–1359, http://dx.doi.org/10.1109/TKDE.2009.191.

[134] J. Tao, T. Han, R. Li, Deep-reinforcement-learning-based intrusion detection in aerial computing networks, IEEE Netw. 35 (4) (2021) 66–72, http://dx.doi.org/10.1109/MNET.011.2100068.

[135] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, Z. Han, Uav-enabled secure communications by multi-agent deep reinforcement learning, IEEE Trans. Veh. Technol. 69 (10) (2020) 11599–11611, http://dx.doi.org/10.1109/TVT.2020.3014788.

[136] Z. Li, Y. Lu, X. Li, Z. Wang, W. Qiao, Y. Liu, Uav networks against multiple maneuvering smart jamming with knowledge-based reinforcement learning, IEEE Internet Things J. 8 (15) (2021) 12289–12310, http://dx.doi.org/10.1109/JIOT.2021.3062659.

[137] C. Wen, Y. Fang, L. Qiu, Securing uav communication based on multi-agent deep reinforcement learning in the presence of smart uav eavesdropper, in: 2022 IEEE Wireless Communications and Networking Conference, WCNC, 2022, pp. 1164–1169, http://dx.doi.org/10.1109/WCNC51071.2022.9771555.

[138] X. Lu, L. Xiao, C. Dai, H. Dai, Uav-aided cellular communications with deep reinforcement learning against jamming, IEEE Wirel. Commun. 27 (4) (2020) 48–53, http://dx.doi.org/10.1109/MWC.001.1900207.

[139] H.v. Hasselt, A. Guez, D. Silver, Deep reinforcement learning with double q-learning, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI '16, AAAI Press, 2016, pp. 2094–2100.

[140] J. Schulman, S. Levine, P. Moritz, M.I. Jordan, P. Abbeel, Trust region policy optimization, 2017, arXiv:1502.05477. https://arxiv.org/abs/1502.05477.

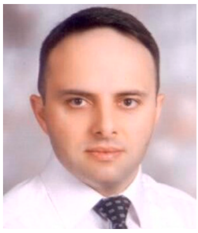
[141] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N.M.O. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, Continuous control with deep reinforcement learning, 2015, CoRR abs/1509.02971. https://api.semanticscholar.org/CorpusID:16326763.

[142] S. Dankwa, W. Zheng, Twin-delayed ddpg: A deep reinforcement learning technique to model a continuous movement of an intelligent robot agent, in: Proceedings of the 3rd International Conference on Vision, Image and Signal Processing, ICVISP 2019, Association for Computing Machinery, New York, NY, USA, 2020, http://dx.doi.org/10.1145/3387168.3387199.

[143] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, I. Mordatch, Multi-agent actor-critic for mixed cooperative-competitive environments, 2020, arXiv:1706.02275, https://arxiv.org/abs/1706.02275.

[144] P.K. Chittoor, B. Chokkalingam, L. Mihet-Popa, A review on uav wireless charging: Fundamentals, applications, charging techniques and standards, IEEE Access 9 (2021) 69235–69266, http://dx.doi.org/10.1109/ACCESS.2021.3077041.

Burcu Sönmez Sarıkaya is a Ph.D. student in the Department of Computer Engineering at Istanbul Technical University. She is a member of Cyber Security and Privacy Research Laboratory, SPF LAB, at Istanbul Technical University. Her research interests include UAVs security, machine learning, deep learning security.



Dr. Şerif Bahtiyar is an associate professor in the Department of Computer Engineering at Istanbul Technical University. He received his BS in Control and Computer Engineering and MS in Computer Engineering degrees both from Istanbul Technical University respectively, and his Ph.D. degree in Computer Engineering from Bogaziçi University. Dr. Bahtiyar was with MasterCard, TU-Berlin in Germany, and National Research Institute of Electronics and Cryptology. Dr. Bahtiyar is the founder and the director of Cyber Security and Privacy Research Laboratory, SPF LAB, at Istanbul Technical University. His current research includes cyber security and privacy, mobile systems, trust modeling, machine learning, e-health, UAV, and financial systems.