Game Theory and Reinforcement Learning for Anti-jamming Defense in Wireless Communications: Current research, Challenges and Solutions

Luliang Jia, Nan Qi, Zhe Su, Feihuang Chu, Shengliang Fang, Ming Xiao, *Senior Member, IEEE,* Mikael Skoglund *Fellow, IEEE,* Kai-Kit Wong, *Fellow, IEEE*, Dobre Octavia,*Fellow, IEEE*, and Chan-Byoung Chae, *Fellow, IEEE*

***Abstract*—Due to the inherent open and shared nature of the wireless channel, wireless communication networks are vulnerable to jamming attacks, and corresponding anti-jamming measures are of utmost importance to realize reliable communication. Game theory and reinforcement learning (RL) are powerful mathematical tools in anti-jamming field. This article investigates the anti-jamming problem from the perspective of game theory and RL. First, different anti-jamming domains and anti-jamming strategies are discussed, and technological challenges are globally analyzed from different perspectives. Second, an in-depth systematic and comprehensive survey of each kind of anti-jamming solutions (i.e., game theory and RL) is presented. To be specific, some game models are discussed for game theory based solutions, including Bayesian anti-jamming game, Stackelberg anti-jamming game, stochastic anti-jamming game, zero-sum anti-jamming game, graphical/hypergraphical anti-jamming game, and so on. For RL-based anti-jamming solutions, some kinds of RL are given, including Q-learning, multi-armed bandit (MAB), deep reinforcement learning and transfer RL. Third, the strengths and limitations are analyzed for each kind of anti-jamming solutions. Finally, we discuss the deep integration of game theory and RL in anti-jamming problem, and some future research directions are presented.**

***Index Terms*—Wireless security, anti-jamming communication, game theory, reinforcement learning, incomplete information, jamming attacks**

I. Introduction

Nowadays, wireless networks have dramatically attracted significant attention, and are highly relevant in both civilian and military applications. However, due to the inherent shared and open nature of the transmission medium, wireless communication is highly vulnerable to various security attacks, such as eavesdropping attacks, spoofing attacks, and jamming attacks. The communication security problem is an important yet challenging concern, and considerable attention has been given in the past decade [1]–[6]. Among these security attacks, the jamming attack is a critical threat, which can deliberately disrupt and deteriorate normal communications [7], [8], and therefore this article focuses on jamming attack.

As a natural consequence, anti-jamming technologies were paid some attention to fight against jamming attacks, and various anti-jamming measures were proposed in existing works [9]–[18]. Spread spectrum-based anti-jamming methods are common to provide anti-jamming capability [19], [20], such as frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS) , and they have been widely adopted in commercial and military applications. Then, considerable efforts have been devoted to improving the performance of spread spectrum-based methods. In [21]–[23], an uncoordinated frequency hopping (UFH) scheme was proposed to cope with the limitation of the pre-shared secrets for traditional spread spectrum technology. In [24], based on the intractable forward decoding and efficient backward decoding, an efficient spread spectrum scheme was designed, and pre-shared secret was not needed. In [25]–[27], a message-driven frequency hopping (MDFH) scheme was investigated to deal with the collision effect of conventional frequency hopping. In [28], a code-controlled MDFH scheme was proposed to pursue better anti-jamming performance, in which block coding technology was integrated into frequency hopping. In [29], adaptive frequency hopping (AFH) mechanism was analyzed, and hopset adaptation scheme was employed. In [30], [31], mode-frequency hopping (MFH) was proposed, in which the angular/mode dimension was considered. In [32], [33], combining the frequency hopping and index modulation, index modulation based FHSS (IM-FHSS) scheme was designed to cope with rective jamming. Although some variants are designed, they have drawn great attention in various scenarios. However, the spread spectrum based anti-jamming solutions imply low spectral efficiency, and need wideband spectrum, which limits their applications with scarce spectrum resource scenarios and band-linmited applications.

The growing wireless devices results in a serious spectrum scarcity problem, and mutual interference among legitimate users is quite prominent. The traditional spread spectrum based anti-jamming approaches are difficult to apply directly in spectrum scarcity scenarios. For these scenarios, optimal resource allocation based anti-jamming methods were promising means [34]–[37], which utilize the spectrum resources in a flexible and efficient way, and constitute another family of anti-jamming measures. What's worse, the jamming technology meets artificial intelligence and cognitive radio (CR) [38], [39], jamming patterns are increasingly complex, and intelligence level is improving. The complex and higher-level intelligent jamming attacks pose severe challenges to the traditional anti-jamming methods. Therefore, the new jamming environment

TABLE I
EXISTING SURVEYS OF ANTI-JAMMING METHODS

Year	Ref.	Key contributions
2009	[9]	A survey of various jamming attacks and typical anti-jamming measures in wireless sensor networks.
2011	[10]	A brief survey of jamming game with incomplete information relying on only 10 citations.
2011	[11]	A survey of jamming models, detection mechanisms and conventional anti-jamming countermeasures in wireless networks.
2014	[12]	A survey of jamming techniques, jamming localization, jamming detection and countermeasures in wireless ad hoc networks.
2016	[13]	A survey of different jamming attacks, jamming detection and defense strategies for wireless local area network, wireless sensor network and ad hoc network.
2020	[14]	A brief survey of dynamic spectrum anti-jamming communication relying on only 15 citations.
2020	[15]	A survey of jamming and anti-jamming techniques in CR network.
2022	[16]	A brief survey of game-theoretic learning anti-jamming relying on 15 citations.
2022	[17]	A brief survey of intelligent dynamic spectrum anti-jamming communication relying on only 15 citations.
2022	[18]	A survey of jamming and anti-jamming schemes from the perspective of different wireless networks scenarios, such as wireless local area networks, cellular networks, vehicular networks, and so on.

will show some typical characteristics, such as higher-level intelligent jammers and complex jamming and interference relationships. Note that the complex jamming and interference relationships mean that the jamming environment suffers from both the mutual interference among users and malicious jamming in this article.

Inspired by the above, it is timely and important to develop effective anti-jamming approaches. Unfortunately, some technical challenges arise concerning the anti-jamming problem, such as incomplete and unknown information constraints, dynamics and confrontation. To address these challenges, some powerful solutions have been investigated. Among these, game theory [40] and reinforcement learning (RL) [41] stand out to analyze and solve the anti-jamming problem, and to design effective anti-jamming schemes in wireless networks. Due to the confrontational charactersitic between legitimate users and jammers, there is a natural interation behavior. Fortunately, game theory, as a well-developed mathematical tool, can adequately formulate the mutual interactions between legitimate users and jammers. Besides, considering the non-cooperative behavior between legitimate users and jammers, incomplete and unknown information constraints are inevitable in dynamic jamming environment. However, RL is an effective method to make sequential decisions in unknown and dynamic environments, which can cope with the incomplete and unknown information constraints and dynamic characteristics. Therefore, game theory and RL are promising methods to achieve better anti-jamming performance in wireless networks.

Fortunately, lots of studies have been devoted to developing effective anti-jamming methods. Preliminary results (i.e., [14]–[16], [42]–[55]) demonstrate the anti-jamming methods based on game theory and RL, which have attained growing attention in anti-jamming problem. Consequently, there is an urgent need for a survey of these anti-jamming methods, where anti-jamming communications techniques meet game theory and RL. Therefore, this article focuses on providing a comprehensive review of the state-of-the-art on anti-jamming methods from the perspective of game theory and RL.

Historically speaking, there are some related surveys on anti-jamming methods. In [9], various jamming attacks were analyzed, and typical anti-jamming measures were investigated in wireless sensor networks. In [10], a brief survey of jamming game with incomplete information was given relying on only 10 citations. In [11], various jamming models were discussed, and some conventional anti-jamming countermeasures were analyzed. In [12], a survey of jamming techniques was presented in detail, and some jamming localization, detection and countermeasure schemes were extensively provided in wireless ad hoc networks. In [13], different jamming attacks, jamming detection and defense strategies were discussed for wireless local area network, wireless sensor network and ad hoc network. In [14], a brief survey of dynmic spectrum anti-jamming communication was presented relying on only 15 citations. In [15], various jamming and anti-jamming techniques were discussed in CR network. In [16], a brief survey of gane-theoretic learning anti-jamming was analyzed relying on only 15 citations, and three kinds of game models (i.e., Stackelberg game, stochastic game and hypergraphical game)

were discussed. In [17], a brief survey of intelligent dynamic spectrum anti-jamming communication was discussed from the persective of deep reinforcement learning (DRL) relying on only 15 citations. In [18], jamming and anti-jamming schemes were analyzed from the perspective of different wireless networks scenarios, such as wireless local area networks, cellular networks, vehicular networks, and so on. The related surveys are summarized in Table I.

In contrast to existing survey papers, this article surveyed the anti-jamming schemes from the perspective of game theory and RL. With respect to the existing survey pspers, the main contributions of this article are given as follows:

(1) The different anti-jamming domains (i.e., power domain, frequency domain, time domain, space domain, code domain and multi-domain) and anti-jamming strategies (i.e., confrontation, avoidance, elimination, hide, tolerance, deceit and bypass) are comprehensively analyzed, and a global analysis of technological challenges is presented from different perspectives in anti-jamming problem, such as network characteristics, information constraints, and the design for anti-jamming decision making scheme.

(2) An in-depth systematic and comprehensive survey of game theory and RL in anti-jamming communications is presented. Specifically, various kinds of anti-jamming game models are discussed for game theory based anti-jamming solutions, including Bayesian anti-jamming game, Stackelberg anti-jamming game, stochastic anti-jamming game, zero-sum anti-jamming game, graphical/hypergraphical anti-jamming game, Colonel Blotto anti-jamming game, anti-jaming relay game, prospect-theory based anti-jamming game and evolutionary anti-jamming game. For RL-based anti-jamming solutions, some kinds of RL are given, including Q-learning, multi-armed bandit (MAB), DRL and transfer RL.

(3) The strengths and limitations are discussed and analyzed for each kind of anti-jamming solution (i.e., game theory and RL). Then, we analyze the integration of game thory and RL in anti-jamming problem, and some future research directions are presented.

The rest of this article is organized as follows. In Section II, background and challenges of anti-jamming communications are investigated. In Section III-V, we mainly analyze and compare two kinds of anti-jamming solutions, respectively. Specifically, game theory based anti-jamming solutions in Section III, RL based anti-jamming solutions in Section IV. In Section V, we discuss the deep intergration of game theory and RL in anti-jamming problem, and some future research directions are discussed. Finally, conclusions are provided in Section VI.

For convenience, the used abbreviations in this article are summarized in Table II

TABLE II
SUMMMARIZATION OF ABBREVIATIONS

Abbreviation	Explanation
AFH	Adaptive frequency hopping
AWGN	Additive white Gaussian noise
CR	Cognitive radio
CNN	Convolutional neural network
DSSS	Direct sequence spread spectrum
DRL	Deep reinforcement learning
ESS	Evolutionary stable strategy
FHSS	Frequency hopping spread spectrum
FCN	Fully connected network
HF	High frequency
IM-FHSS	index modulation based frequency hopping spread spectrum
IRS	Intelligent reflecting surface
IoT	Internet of Things
IID	Independent and inentically distributed
LSTM	Long short term memory
MDFH	Message-driven frequency hopping
MFH	Mode-frequency hopping
MAB	Multi-armed bandit
MIMO	Multiple-input multiple-output
MAC	Medium access control
MDP	Markov decision process
NOMA	Non-orthogonal multiple access
NE	Nash equilibrium
OFDM	orthogonal frequency divison multiplexing
PT	Prospect theory
RL	Reinforcement learning
RNN	Recurrent neural network
RIS	Reconfigurable intelligent surface
SINR	Signal to interference plus noise ratio
SE	Stackelberg equilibrium
THSS	Time hopping spread spectrum
UFH	Uncoordinated frequency hopping
UAV	Unmanned aerial vehicle

II. Background and Challenges of Anti-Jamming Communications

A. Discussion of jammers

To better solve the anti-jamming communication problem, it is necessary to understand communication jamming. Jamming is a deliberate and intentional emission of wireless signals, which aims to disrupt or prevent normal communication [9]–[15], [56]–[58], and comes from external jammers or malicious users. According to different classification rules, jammers can be classified into different types. Based on action modes of jammers, they can be divided into two categories: suppressive jammer and deceptive jammer. Suppressive jammer, e.g. constant jammer and reactive jammer, refers to the transmission of jamming signals in communication frequency band, which deteriorates the signal to interference plus noise ratio (SINR) of receiver, and reduces or loses the ability of receiver to receive information. Deceptive jammer is to imitate the characteristics of legitimate communication behaviors, so as to deceive legitimate users and mislead them to make unexpected response [9]–[15], [56]–[59]. Based on the functionality, jammers can be divided into elementary jammer (e.g. constant jammer and random jammer) and advanced jammer (e.g. follower-on jammer and control channel jammer) [12]. The class of elementary jammers includes the proactive jammer and reactive jammer. A proactive jammer transmits a jamming signal regardless of data transmission, such as constant jammer and random jammer. The reactive jammer has the sensing ability, and it transmits jamming signal only when the communication activities exist on the channel. For the advanced jammers, they classify as function-specific jammer and smart-hybrid jammer. The function-specific jammers have a pre-determined function to generate a jamming signal, such as follow-on jammer and channel hopping jammer. The smart-hybrid jammers hold effective jamming characteristic, and aim to improve jamming effect to conserve energy, such as control channel jammer and implicit jammer. Based on the level of sophistication, jammers can be divided into smart jammer and non-smart jammer [15]. For the non-smart jammers, they follow a fixed jamming strategy, and interested readers can refer to literatures for detailed information [9]–[12], [56]–[58]. This subsection focuses on the smart or intelligent jammers.

To pursue better jamming effect, smart or intelligent jammer is more effective. In [60], [61], an intelligent adversary, which can be regarded as an combination of spoofing and jamming, was investigated in CR, and it optimized the spoofing signal in sensing duration and optimized the jamming signal in transmission duration. In [62], the optimal jamming problem was analyzed in additive white Gaussian noise (AWGN) channel, and the optimal jamming signal was designed for digital amplitude-phase-modulated constellations. In [63], the optimal jamming signal design was considered in multiple-input multiple-output (MIMO) communication system. In [64], [65], a cognitive jammer based on a MAB framework, which can optimize the physical layer parameters, was designed to optimally jam the communication between one transmitter and receiver. In [66], [67], the intelligent jammer based on RL was investigated to increase the jamming impact. In [68], [69], the intelligent jammer based on DRL was discussed, and a deep learning guided jamming was proposed in [70]. According to the references [15], [71]–[73], the intelligent jammer can learn transmission pattern and dynamically adapt to jamming environment. It should be noted that learning ability and adaptability are the most significant characteristics of intelligent jammer, and it will be a focus of anti-jamming problem in the future work.

B. Discussion of anti-jamming problem

Anti-jamming communication is a classical yet interesting problem, and it plays a fundamental role in supporting the research and development of reliable communication techniques. Moreover, it is also an important part of cognitive risk control [74]–[76]. Anti-jamming communication is a highly topic with several sub-problems. Specifically, it includes jammer detection [77]–[85], jammer recognition [86]–[89], jammer localization [90], and anti-jamming decision making [42]–[55], and so on. An illustration of anti-jamming technology is shown in Fig. 1.

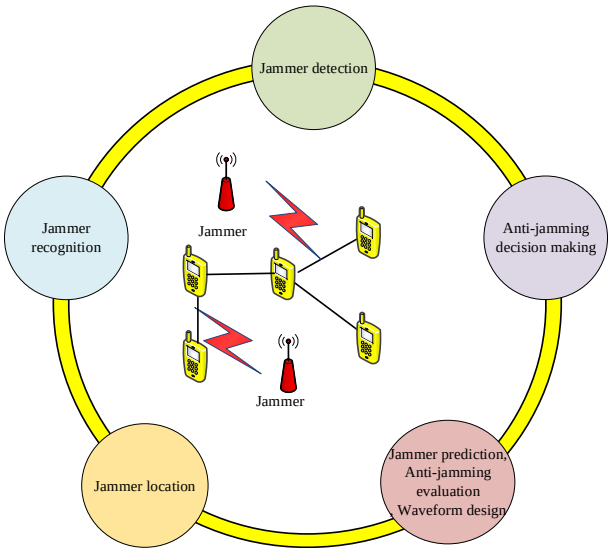

Fig. 1. An illustration of anti-jamming communication problem.

- **Jammer detection** :Jamming detection is the first step of an anti-jamming scheme, and various anti-jamming methods can be performed only when jammer is correctly detected. For the jamming detection problem, there exist some works, such as hypothesis testing [81], compressed sensing [82], and machine learning based method [83], [84]. These methods can be mainly classified into two categories [85]: threshold based methods and medium access control (MAC) based methods. In the threshold-based detection methods, a threshold is employed to differentiate between jamming scenario and legitimate scenario, and they focus on communication between two legitimate users. The MAC-based methods can observe medium access process with predefined duration, and they can distinguish the medium access contention collision and jamming attack.
- **Jammer recognition** : The identification of jamming patterns is another important problem, and it devoted

to understanding and distinguishing different jamming patterns. Because it may need different anti-jamming approaches for different jamming patterns, it is highly meaningful to employ targeted anti-jamming schemes to improve anti-jamming capability. In [86], Naive Bayes classifier was employed, a data driven jamming recognition method was proposed. In [87], a spectrum waterfall is adopted, and a deep learning based jamming recognition method is investigated. In [88], based on a combination of threat-based schemes and nonparametric estimation, an algorithm was designed to estimate jammer type. In [89], based on a distributed recognition framework, a few-shot learning method was proposed for jamming recognition.

- **Jammer localization** : The localization of jamming attacks plays an essential role in designing jamming-resistant countermeasures. Some jamming parameters need to be estimated for anti-jamming measures, such as jamming channel gain and jamming powe, while accurate jamming patermeters may benefit from accurate jammer localization. In [90], a comprehensive survey of jammer localization was presented, and various jamming locailization schemes were provided. The jammer localization problem involves additional concepts, such as information measurement, location computation and accuracy evaluation. Based on the used information of location computation, the existing jammer localization methods can be divided into range-based and range-free schemes. In a range-based method, distance information is employed to obtain a jammer's location. For a range-free method, the network topology information is utilized to locate a jammer.
- **Anti-jamming decision making** : To obtain desirable anti-jamming strategies, such as transmission power and channel, it is of critical importance to design effective anti-jamming decision making approaches in jamming environments. In this article, we focus on anti-jamming decision making, which is a core and critical technology in anti-jamming problem. Meanwhile, it is a challenging task due to the adversarial characteristic and non-cooperative relationship between the legitimate users and malicious jammer. In existing works, there are various anti-jamming decision making schemes, such as game theory based methods [42]–[51], RL based methods [52]–[55], convex optimization based methods [91]–[99], and particle swarm optimization based methods [100]. Among these methods, game theory based methods and RL based methods have some advantages in dynamic, incomplete and unknown jamming environment, and we will discuss later. This survey mainly focuses on the game theory based methods and RL based methods.

Remark 1. *Besides the game theory based methods and RL-based methods, there are also some other methods, e.g., convex optimization based methods and particle swarm optimization based methods. However, these methods have high requirements for jamming environment information, and they are suitable for scenarios with known and complete environmental information. In this survey, we do not discuss these methods, and they belong to the class of centralized defense schemes and may incur significant communication cost due to parameter estimation and result in latency in the network.*

In addition, the anti-jamming problem involves additional contents, such as jammer prediction [101], anti-jamming evaluation [102], [103] and waveforms design [104]. Based on jammer prediction, the jamming rules can be learned according to the jammer's historical information, and legitimate users can avoid jammers' behaviors in advance. Through anti-jamming performance evaluation, the weakness of anti-jamming methods can be found, and then it is beneficial to further improve the designed anti-jamming schemes. The design of waveform is physical countermeasure, and the suitable anti-jamming transmission waveforms need to be designed to enhance the anti-jamming performance in wireless networks.

C. Discussion of anti-jamming domains

From the perspective of anti-jamming domains, different domains are employed to combat the impact of jamming attacks, such as power domain [105]–[111], frequency domain [19]–[32], [112]–[116], time domain [104], [117]–[122], space domain [123]–[131], code domain [11], [132]–[135] and multi-domain [136]–[144]. Correspondingly, the anti-jamming approaches can be categorized into the following categories.

- **Power domain** : The power domain anti-jamming methods are to combat jammers at the cost of power, and higher power means stronger resistance to jamming attacks. When jamming power is not very strong, as long as it can meet the normal communication requirements, the smaller the transmitting power, the better. In case of high-power strong jammer, it is the most direct way to adopt high-power to fight against jamming attack. With the development of miniaturization and intelligence of jamming attacks, the coexistence of high-power suppression jammers and low-power smart jammers will appear in the future.
- **Spectrum domain** : The spectrum domain anti-jamming methods are mainstream anti-jamming methods, and have attracted widespread attention. The spread spectrum based anti-jamming schemes can achieve anti-jamming ability by spreading the signal bandwidth to a wider frequency band, such as FHSS and DSSS, and the reliability requirements can be meet at the at the expense of spectrum efficiency. Moreover, legitimate users can avoid the jamming attacks by looking for or employing spectral holes. Besides, some effective anti-jamming schemes are proposed in spectrum domain, such as dynamic spectrum anti-jamming [14] and game-theoretic learning anti-jamming paradigm [16], which can take full advantage of spectrum resources and adapt to spectrum environment.
- **Time domain** : The transmission time can be subdivided into time slots, and the information transmission can be realized according to the jamming time interval. Meanwhile, time hopping spread spectrum (THSS) is a typical anti-jamming scheme, where the signal shifts transmission slots in a pseudorandom way, which has

some advantages, such as low implementation complexity, and low interception probability [104]. In addition, the secondary users do not have spectrum-access priority in CR, and spectrum-based measures may be ineffective in jamming environments due to higher channel switching rate [117], [118]. For these scenarios, time-based anti-jamming schemes may be a good candidate. Based on RL, a time domain anti-jamming pulse jamming mechanism was proposed in [119]. In [120], the spreading-time technology was employed for low power and band-limited nodes in Internet of Things (IoT), and an automatic control allocation framework was formulated. In [121] , a novel time hopping anti-jamming scheme was proposed, and the non-ccoherent chaotic system was employed.

- **Space domain** : Space domain based anti-jamming schemes exploit the spatial dimension to cancel out the impact of jamming attacks based on diversity of space domain and network topology characteristics, and can be applied in multi-antenna communication systems. Based on multi-antenna techniques, legitimate users can cope with jamming signals with antenna gain. Different from frequency hopping based defense schemes, space domain based anti-jamming schemes can recover the normal signal in jammed spectrum, which can realize the jamming defense through spatial freedom. Based on the diversity and multiplexing gain, MIMO-based anti-jamming schemes were proposed in [123]–[126], and anti-jamming capability can be achieved by spatial filtering. In [127], [128], the anti-jamming scheme was explored by controlled mobility, and the legitimate nodes move their geographical locations to fight against jamming attacks. In [129]–[131], based on the spatial diversity of relay nodes, cooperative anti-jamming relaying schemes were proposed in vehicular networks.
- **Code domain** : In the code domain, suitable coding and modulation schemes can be employed to the robustness for jamming attacks. For example, error-correction codes, such as Low Density Parity Codes and Turbo-codes, aim to enhance the error tolerance. To be specific, the transmitter encodes the information by adding redundant bits, and the receiver obtains the anti-jamming ability by decoding the transmitted information. Then, some error bits can be recovered with extra transmission bits. From the perspective of jammers, only a small amount of bits need to be disrupted to cause a transmission failure in scenarios without error-correction, while higher jamming price are needed to corrupt the transmission for error-correction scenarios. Moreover, if the received SINR is lower than the demodulation threshold, low rate and reliable modulation is suitable, e.g., binary phase shift keying.
- **Multi-domain** : Multi-domain defense mechanisms mean that a variety of anti-jamming strategies are flexibly adopted in multiple domains to obtain better anti-jamming performance. In [136]–[142], frequency hopping and power adaptation were jointly considered. In [143], a multi-domain anti-jamming scheme was provided, which combined the time domain and spectrum domain to design the jamming countermeasures. In [144], a joint time-frequency jamming-resistant scheme was investigated, and a reinforcement learning based algorithm was proposed to obtain the optimal channel and transmission duration.

D. Discussion of anti-jamming strategies

According to the different anti-jamming strategies, anti-jamming methods can be mainly divided into seven categories [14], [16], [145]: confrontation, avoidance, elimination, hide, tolerance, deceit, and bypass.

- **Confrontation** : "Confrontation" means that legitimate users can directly increase the transmitting power to combat jamming attacks, and therefore it results in the increase of the received SINR. For example, power control anti-jamming employs this strategy to cope with the threat of jamming attacks in [105]–[110].
- **Avoidance** : This anti-jamming strategy aims to escape jammers. For example, frequency hopping switches between different channels when the current channel is jammed. In addition, legitimate users can avoid the direction of incoming wave, and form the spatial isolation between communication signal and jamming signal to reduce the impact caused by jammers. Finally, spatial retreat is another effective measure, which can move on others locations when the current area suffers from heavy jamming [146].
- **Elimination** : Different from other anti-jamming strategies, "elimination" means that the jamming in the received signal can be eliminated as much as possible by means of signal processing, such as adaptive filtering and blind source separation.
- **Hide** : "Hide" seeks to improve the concealment of the communication signal so that the jammers cannot detect the existence of communication signal. For example, DSSS submerges the communication signals in noise for transmission. In addition, stealthy communication is a reliable and stealthy communication paradigm, and communication behavior is undetectable by adversary. Meanwhile, the normal communication should ensure robustness to jamming attacks [147].
- **Tolerance** : By "tolerance", some coding schemes can be adopted to disperse and reduce the impact of jamming, such as error-correction code and repeated transmission. The jamming defense ability is obtained by redundancy information in error-correction codes.
- **Deceit** : "Deceit" is an important anti-jamming strategy, and it can realize the anti-jamming communication by fake information to lure jammers. In [148], an extra transmitter-receiver pair was added to send fake information to attract the jammer to ensure the real information transmission. In [149], [150], the transmitter deliberately takes some wrong actions to fool the jamming attack into making prediction errors, and the anti-jamming performance can be enhanced. In [151]–[155], various deception strategies were proposed to combat reactive jammers. In [156]–[159], the deception tactics were investigated to cope with intelligent jamming attacks.

- **Bypass** : Different from conventional physical-layer anti-jamming techniques, (e.g., spread spectrum-based methods) and MAC-layer schemes (e.g., error-correction), "bypass" employs a different perspective, and designs anti-jamming mechanisms at network level. It resorts to making full use of routing diversity, and re-establishing routing connectivity. To combat the jamming attacks, bypassing mechanisms can be employed to detour the jammer-affected zone, and a new alternative paths can be established. In [145], [160]–[166], routing approaches were investigated to bypass the jammers, and various effective path selection algorithms were proposed.

E. Discussion of Challenges

In this subsection, the main challenges of the anti-jamming decision making problem is summarized from different perspectives [14], [16]. The main challenges facing anti-jamming decision making and the corresponding approaches are summarized in Fig.2.

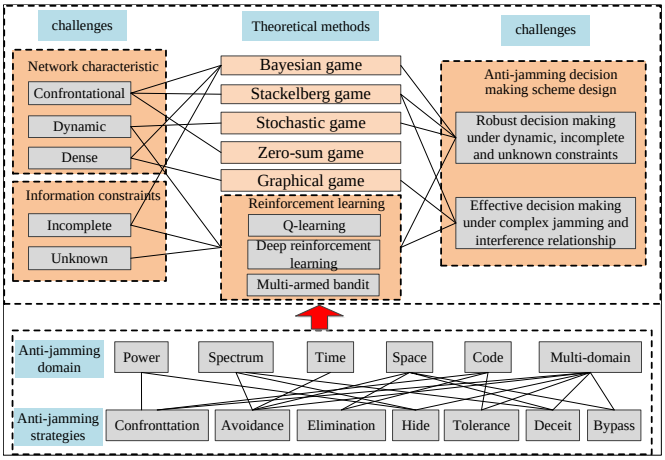

Fig. 2. The challenges and solutions in anti-jamming problem.

First, from the perspective of network characteristics, three challenges need to be addressed in anti-jamming field.

- **Confrontational** : It consists of two classes of main entities in the anti-jamming problem: the legitimate users, who pursue reliable communication in adversarial jamming environments; and malicious jammers aims to deteriorate the normal information transmission. The confrontation is the most obvious characteristic for anti-jamming problem. In addition, with the development of intelligent jamming technology, the intelligent jammer can learn the transmission pattern and adapt to the dynamic strategy of legitimate users, and the confrontation process is more intense. This will further increase the difficulty for designing of jamming defense scheme, and "intelligent anti-intelligent jammer" will be a severe challenge for anti-jamming communication in the future. It should be pointed out that equivalent countermeasure capability is the premise for legitimate users and jammers to confront each other, and they realize the optimization of their own utilities through adjusting their own strategies. If one side has absolute advantages for legitimate users and jammers, the weak side does not have any effective countermeasures.
- **Dynamic** : The dynamic is very important, and there are several factors. First, the spectrum environment is dynamic in jamming environment duo to adversarial interactions between legitimate users and malicious jammers. Second, jammer environment may be dynamic, such as jamming frequency, jamming power, and jamming location. In addition, the dynamic characteristic of wireless communication network is also important. For example, spectrum state is time-varying, and traffic demands may change from time to time as well.
- **Dense** : With the continuous growth of traffic demands, wireless devices are deployed intensively, such as UAV networks and small cell networks. It is a challenging task to develop jamming defense schemes in dense networks. First, malicious jamming is a potential threat to degrade the information transmission of the legitimate users. Second, mutual interference, which brings among legitimate users with the same channel, is another important factor to restrict the network performance. Therefore, the jamming environment becomes complex in dense networks, and spectrum confrontation is very fierce.

To realize reliable communication, both malicious jamming and mutual interference need to be considered. For this, there exist some studies in the existing literature. In [167], [168], both the inter-cell-interference and malicious jamming were investigated in heterogeneous cellular networks, and decoupled association and reverse frequency allocation were exploited to combat both jamming and interference. In [169], a SimpleMAC protocol was formulated, and a channel coordination mechanism was designed to minimize interference among transmitters, and yet the jammer can be prevented. In [170]–[172], anti-jamming schemes were discussed in interference alignment networks, and jamming signal and interference alignment are considered jointly. In [173], the anti-jamming communication issue was investigated in the ultra-dense network, and frequency-hopping technology was adopted to cope with jamming and interference. In [174], non-orthogonal multiple access (NOMA) and distributed antenna system was combined to counteract the jamming, and simultaneously mutual interference can be removed.

Second, from the perspective of information constraints, the following challenges arise in anti-jamming problem.

- **Incomplete** : The design of jamming defense mechanisms needs some information. Unfortunately, it is impossible for legitimate users and jammers to obtain the perfect and accurate information of each other due to the non-cooperative relationship. Besides, owing to the limitation of signal processing capability and resource, each entity for both legitimate users and jammers can only obtain part information of wireless environment. Therefore, incomplete information constraint is unavoidable in wireless jamming environment. In existing studies, various kinds of incomplete information needs to be analyzed.

  In [9], various forms of incomplete information constraint

were analyzed, such as user types, system parameters and physical channel, and Bayesian game framework was presented to deal with the incomplete information constraints. In [175], [176], the rival-type uncertainty was analyzed, and various Bayesian anti-jamming game models were provided. In [177], a Bayesian power control anti-jamming game was formulated, and the incomplete nature of channel gain was analyzed. In [178], [179], the position uncertainty was investigated, and the corresponding Bayesian game was formulated.

- **Unknown** : In some anti-jamming scenarios, partial information can be obtained. Unfortunately, due to the adversarial characteristic, the information of confrontation environment is completely unknown in some cases, such as jammer pattern, and jamming parameters, and prior information of jamming environment cannot be achieved. It is a challenging task to design an effective anti-jamming scheme in unknown jamming environments.

It should be noted that there exists some research addressing anti-jamming problem in unknown jamming environment. In [52], [53], Q-learning, as a typical reinforcement learning method, was employed to overcome the impact of jammers in unknown jamming environment. In [14]–[17], [173], [180], [181], the deep reinforcement learning based anti-jamming methods were investigated in various unknown jamming environments. In [141] and [182], based on the MAB framework, and online learning algorithms were proposed to obtain the desirable anti-jamming strategies in unknown environment.

Third, from the perspective of the design for anti-jamming decision making schemes, two challenges are listed as follows.

1) **Robust decision making under dynamic, incomplete and unknown constraints**. Due to the adversarial characteristic and inherent features of the jamming environment, the dynamic, incomplete and unknown information constraints need to be addressed in anti-jamming decision making. The Bayesian game framework is an alternative method to cope with the incomplete information constraint, and it optimizes the expected utility function of game player depends only on distribution information rather than accurate information. In addition, intelligent learning methods, such as RL and DRL, are effective to cope with the dynamic, incomplete and unknown information constraints, and they can constantly interact with the environment through trial and error, so as to achieve the best match between their strategies and the jamming environment.

2) **Effective decision making under complex jamming and interference relationship**. In some anti-jamming scenarios, the effects of mutual interference and malicious jamming need to be counteracted at the same time, and spectrum confrontation is fierce. It is a challenging task to design effective anti-jamming decision making methods under complex jamming and interference relationship scenarios. Therefore, two basic ideas can be employed. First, mutual interference and malicious jamming can be mapped into generalized interference and jamming through a certain functional relationship, and then the minimization problem of generalized interference and jamming can be solved. Combined with intelligent learning methods, the generalized interference and jamming problem was discussed in [183], [184]. Second, the idea of internal collaboration and external confrontation can be exploited, and both collaboration and competition were considered. To be specific, by collaboration mechanisms among legitimate users, the mutual interference can be eliminated while avoiding malicious jamming in the process of learning jamming behaviors. In [185]–[188], the collaboration among legitimate users was considered, and collaborative learning algorithms were proposed to simultaneously tackle the mutual interference and malicious jamming.

III. Game theory-based anti-jamming approaches

Game theory, as a powerful mathematical tool, can describe and analyze competitive/cooperative interactions in multi-player scenarios [40]. According to whether a binding agreement can be formed among players, game theory can be divided into two branches: cooperative and non-cooperative. In the case of cooperative model, game players make decisions based on cooperative means, and there are certain enforceable agreements among players. Therefore, one player considers the impact for other players in the process of making decisions. For a non-cooperative game, each game player makes rational decisions independently and selfishly without considering the impact of its own behavior for other game players, and aims to maximize its individual utility function. In addition, game theory can be categorized into complete information game and incomplete information game according to the available information of players. Each player has accurate information of other players in a complete information game, such as strategy space and utility functions, while only partial information can be available in an incomplete information game. In anti-jamming field, the non-cooperative and incomplete information game are mainly considered.

In this section, the fundamentals of game theory are firstly provided, and then some state-of-the-art anti-jamming game-theoretic solutions are analyzed and reviewed. Finally, strengths and limitations are discussed.

A. Fundamentals of game theory

A non-cooperative game consists of three elements: player, strategy set and utility function, and a strategic form game can be expressed as a triplet $\mathcal{F} = \{\mathcal{N}, \mathcal{A}_n, \mu_n\}$, where $\mathcal{N}$ is the set of game players, $\mathcal{A}_n$ represents the strategy set of game player n, and μ_n denotes the utility function.

- **Player** : Entities, who can independently make rational decisions, are called players. Players have clear goals, and more than one action is available. The player set is a finite set, and it can be denoted by $\mathcal{N} = \{1, ..., n, ..., N\}$. A game has at least two players, and both legitimate users and malicious jammers can act as players in anti-jamming problem.

TABLE III
SUMMARY OF GAME MODELS IN ANTI-JAMMING FIELD

Game model	Advantages	Utility function	Techniques	Ref.
Bayesian game	It can cope with the incomplete information constraints.	Throughput; successful transmission rate	Convex optimization techniques	[10]
		Throughput	Convex optimization techniques	[175], [189], [190], [191], [251]
		The number of successful receptions	Convex optimization techniques	[177], [193]
		Channel capacity	Linear programs	[178], [179]
		SINR	Convex optimization techniques	[194], [195], [196], [190], [191], [197]–[200]
Stackelberg game	It can capture hierarchical behavior, and analyze competitive interactions at different levels.	SINR	Convex optimization techniques	[105], [106], [191], [199], [200], [201]–[209], [210]–[212]
		SINR	Q-learning	[213], [214]
		SINR	Genetic algorithm	[143]
		SINR	Deep Q networks	[215]
		Throughput	Convex optimization	[148], [216], [217], [218]
		Weighted aggregate interference and jamming	Stochastic learning	[183]
		Weighted aggregate interference and jamming	Log-linear learning	[219]
		The number of successful receptions	Q-learning	[220]
		Throughput	Deep neutral network	[221]
		Throughput	Distributing learning	[222]
		Throughput	Better reply algorithm	[223]
		Throughput	Q-learning	[224]
		Energy consumption	Q-learning	[225]
		Multi-objective cost	Deep reinforcement learning	[226], [227]
		The number of useful communication	Linear programs	[228]
Stochastic game	It can describe the dynamics of jamming environment.	The number of successful receptions	Multi-agent reinforcement learning	[185]–[188]
		Throughput	Deep reinforcement learning	[229], [230]
		Throughput	Minimax-Q learning	[231]
		Throughput	Multi-agent reinforcement learning	[232], [233], [234], [235]
		Transmission rate	Linear programs	[236], [237]
		Throughput	Linear programs	[238]
Zero-sum game	It can capture the adversarial relationship between legitimate users and jammers.	Capacity	Convex optimization	[239], [240]
		Capacity	Q-learning	[241]
		Throughput	Linear program	[242]
Graphical/ hypergraphical game	It can well model the mutual interference effect among legitimate users.	Throughput	Distributed learning	[243]
		Satisfaction	Distributed learning	[244]
		Generalized interference and jamming	Distributed learning	[184]

- **Strategy** : A strategy can be regarded as a mapping from available information to actions. Actually, it is also known as a possible decision made by a player based on the available information. Strategies can be divided into pure strategies and mixed strategies. The chosen strategy of player n can be expressed as $a_n \in \mathcal{A}_n$, and a strategy profile of all players can be denoted as $a = \{a_1, ..., a_N\}$. A player chooses a unique action from its action set in a pure strategy, while a player randomly selects a strategy in strategy set according to a certain probability distribution in a mixed strategy. Specifically, a mixed strategy of player n can be represented by $\theta_n(a_n)$, which means the probability that player n chooses strategy $a_n \in \mathcal{A}_n$ and satisfies $\sum_{a_n \in \mathcal{A}_n} \theta_n(a_n) = 1$ and $0 \leq \theta_n(a_n) \leq 1$. Then, a mixed strategy profile of all players can be denoted as $\theta = \{\theta_1, ..., \theta_N\}$.
- **Utility function** : The utility function describes the clear goal of a player, and expresses the preferences for different strategies. It can also be regarded as a reward obtained by players. For a non-cooperative game, each player aims to maximize its own utility function, and a player can obtain their desirable strategies by maximizing its utility function, which is driven by self-interest. It should be pointed out that the utility function is closely related to the properties of the game model. For the continuous optimization scenarios, if the utility function is convex, the traditional convex optimization methods can be applied to achieve the optimal strategy for the continuous problems. For the discrete optimization scenarios, the design of utility function is also very important. For example, the exact potential game can be used in anti-jamming channel selection problem [16], and it satisfies the following constraints that the variation of the utility function due to any game player's unilateral deviation is equal to the variation of the potential function. For an exact potential game, it has at least one pure strategy Nash equilibrium (NE), and any global or local maximization of potential function can constitute an optimal pure strategy NE. In addition, the ultimate goal of formulated anti-jamming game is to obtain the desirable anti-jamming strategies, and the utility function should have clear physical meaning, such as SINR and throughput.

Another important terminology of the non-cooperative game is NE [40], which is the basic solution concept. In order to analyze the mutual interactions among players and the results of self-interest, NE is the common steady solution of a non-cooperative game. In a NE, no player can obtain the improvement of utility function to unilaterally change its strategy. Specifically, a strategy profile $a^* = \{a_1^*, ..., a_N^*\}$ is a pure strategy NE if no player has an incentive to unilaterally change current strategy to improve its utility function, and the strategies of other players are fixed. Correspondingly, a mixed strategy profile $\theta^* = \{\theta_1^*, ..., \theta_N^*\}$ is a mixed strategy NE if no player has an incentive by unilaterally changing its current mixed strategy to another strategy. It should be noted that pure strategy NE can be regarded as a special case of mixed strategy NE, in which each player choose one of strategies with probability 1, and the probability of selecting other strategies is zero.

B. The applications of game theory in anti-jamming communications

Nowadays, game theory provides a rich framework to cope with jamming attacks, and various game models have been employed, such as Bayesian game [175]–[179], Stochastic game [185]–[188], Stackelberg game [245], and so on. Different game models have different characteristics, and the existing game models are reviewed and compared in anti-jamming field in Table III.

1) **Bayesian anti-jamming game**

Due to the non-cooperative and adversarial relationships between legitimate users and malicious jammers, it is necessary to deal with incomplete information constraints. Bayesian game provides a framework to describe and analyze the incomplete information constraints, and it only needs distribution information rather than accurate information. The uncertainty of incomplete information is described by a prior distribution, and then players aim to optimize their expected utility function. The existing Bayesian anti-jamming game applications are reviewed in Table IV.

TABLE IV
EXISTING APPLICATIONS OF BAYESIAN ANTI-JAMMING GAME

Kinds of uncertainty	Forms of uncertainty	Ref.
Player type uncertainty	Rival-type uncertainty	[175], [194], [246], [247]
	Jamming attack type uncertainty	[176]
	User type	[10], [189]
System parameters	Channel gain	[10], [177], [190], [191], [197]–[200]
	Rendezvous channel	[192]
	Jamming power	[195]
Other forms	Position uncertainty	[178], [179], [193]
	Traffic uncertainty	[10], [249], [250]
	Physical presence	[10], [196], [251]

Various forms of uncertainty were discussed, and corresponding Bayesian anti-jamming game models were formulated. In [175], [194], incomplete information of rival's identity is analyzed, and player cannot confirm whether the rival is regular-type or smart-type. Then, the Bayesian anti-jamming game framework was employed to cope with rival-type uncertainty. In [246], the adversary's behavior patterns

are uncertain, and it is not sure whether the adversary will adopt Nash behavior or Stackelberg behavior. In [247], the uncertainty of the rival's activity was analyzed, and the Bayesian game framework was discussed. In [176], the type of intelligence of jamming attacks is incomplete. To be specific, it is uncertain whether the jamming attack is a random jammer or an intelligent jammer. In [189], the uncertainty of user type was analyzed in medium access control game, and each transmitter (i.e., selfish and malicious transmitter) had incomplete information of types for other transmitters.

Due to the adversarial relationships, the uncertainty of system parameters is common. In [177], [197], the incomplete information of channel gains was considered. Specifically, neither the legitimate user nor the jammer known the exact channel gain information of opponent link, and only the probability density functions of random variables were available. In [198], the uncertainty of channel gain was analyzed in an orthogonal frequency divison multiplexing (OFDM) wireless network, and the optimal strategies were obtained based on Bayesian anti-jamming game framework. In [190] , the legitimate user cannot obtain the exact location of jammer in a single carrier wireless system, and only its probabilistic term was known. In [191], [199], [200], the incomplete channel gain was modeled and analyzed in Bayesian Stackelberg anti-jamming game. In [192], a Bayesian channel selection game is proposed in dynamic spectrum access networks in the scenario of unknown rendezvous channel. In [195], the precise jamming power is unknown, and only a statistical description was available.

In addition, there are other forms of uncertainty. In some wireless networks, the performance is position-dependent. In [178], the exact position of legitimate transmitter was unknown for jammer, and only probability distribution was available. In [179], the position uncertainty was analyzed in underwater sensor networks. In [193], considering the incomplete information of the jammer location in a CR network, a Bayesian anti-jamming power control game was formulated. In [249], [250], the traffic uncertainty was considered, and corresponding countermeasures were analyzed. In [196], [251], the uncertainty of physical presence or absence of jamming attack was investigated, and the legitimate user only had the statistical probability that the jammer was present or absent. In [10], Bayesian jamming game framework was discussed, and incomplete information types are analyzed.

2) **Stackelberganti-jamming game**

Stackelberg game, as an extension of the non-cooperative game, has many favourable characteristics. First, it includes two kinds of players with different attributes, namely leaders and followers. Second, it can describe these hierarchical interactions between leaders and followers. Specifically, the leaders have strong position, and they play their strategies first. Then, the followers consequently react to the declared strategies of the leaders. A Stackelberg anti-jamming game can be expressed as $\mathcal{F} = \{\mathcal{N}, \mathcal{J}, \mathcal{A}_u, \mathcal{A}_j, \mu_u, \mu_j\}$, where $\mathcal{N}$ is the set of legitimate users, $\mathcal{J}$ is the set of malicious jammers, $\mathcal{A}_u$ and $\mathcal{A}_j$ respectively represent the strategy space of legitimate users and jammers,μ_u and μ_j are utility function of legitimate users and jammers, respectively. In some anti-jamming scenarios, there are hierarchical behaviors between the legitimate users and jammers. For example, the jammer can learn the transmission strategies of the legitimate users [105], [106], and it is necessary for the legitimate users to detect the jamming actions [245]. Therefore, Stackelberg game is a natural tool to capture these hierarchical interactions. Besides, Stackelberg game can describe multiple competitive relationships. On the one hand, there are hierarchical competitive interactions between the leaders and followers. On the other hand, there are also competitive interactions among followers or leaders. A framework of Stackelberg anti-jamming game is shown in Fig. 3. The existing Stackelberg anti-jamming game applications are reviewed in Table V.

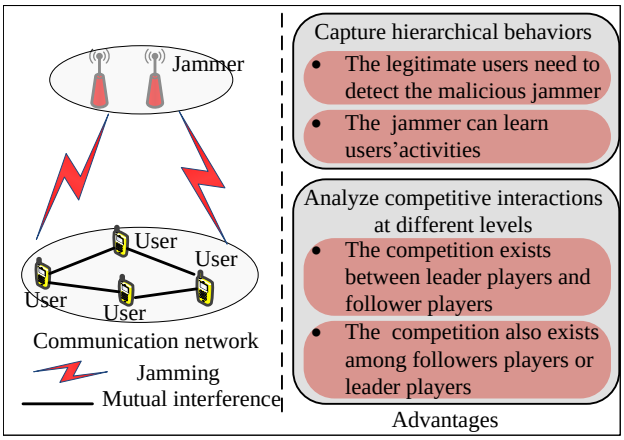

Fig. 3. A framework of Stackelberg anti-jamming game.

Stackelberg game can be regarded as a two-level optimization problem, and it can be applied to solve the anti-jamming decision-making problem with continuous strategy. Based on convex optimization techniques, the Stackelberg equilibrium (SE) can be obtained by a backward induction method. In [105], [106], [148], [216], [202], the one-leader one-follower Stackelberg anti-jamming power control game was formulated, and the desirable power strategies were obtained. In [199] and [203], Bayesian Stackelberg power control game was formulated, and the incomplete information was considered. Based on successive convex approximation method, a confrontation game was formulated to describe the scenario with simultaneous jamming and eavesdropping in [204], and the desirable power solution was achieved. In [205], the observation error of the jammer was investigated in a Stackelberg power control game. In [206], [207], an anti-jamming power control game was analyzed in wireless cyber-physical system. In [208], a jamming mitigation scheme was proposed in power domain for IoT communications. In [209], [256], a jamming avoidance problem was investigated in a wireless sensor network. In [252], beam-domain anti-jamming problem was analyzed in a massive MIMO system, and optimal closed-form power solution was obtained for both the base station and jammer. In [259], based on the ability of constructing wireless environment, a reconfigurable intelligent surface-assisted Stackelberg anti-jamming scheme was presented to improve the anti-jamming performance. In [257], a Stackelberg game was employed to model the power silence to fight against the jamming attacks. In [228], [258],

TABLE V
EXISTING APPLICATIONS OF STACKELBERG ANTI-JAMMING GAME

Application	Strategy space	Situations	Ref.
Power control	Continuous strategy	One-leader one-follower	[105], [106], [148], [191]–[193], [195], [196], [199]–[209], [216], [249]–[252], [259],
	Continuous strategy	One-leader multi-follower	[191], [210], [217], [253]
	Continuous strategy	Three-layer Stackelberg game	[200], [211], [212]
	Discrete strategy	One-leader one-follower	[213], [214], [220], [221], [254], [255]
Mobility-based jamming avoidance	Continuous strategy	One-leader multi-follower	[256]
Power silence	Continuous strategy	One-leader one-follower	[257]
Offloading application	Continuous strategy	Three-layer Stackelberg game	[218]
Channel selection	Discrete strategy	One-leader multi-follower	[183], [245]
	Discrete strategy	Multi-leader multi-follower	[219]
	Discrete strategy	Multi-leader one-follower	[222]
The joint channel selection and power control	Discrete strategy	Multi-leader one-follower	[223]
The joint selection of frequency hopping speed and power control	Discrete strategy	One-leader one-follower	[143]
The formulation of anti-jamming subnetwork formulation	Discrete strategy	Multi-leader multi-follower	[225]
Routing selection	Discrete strategy	One-leader one-follower	[226], [227]
Trajectory optimization	Discrete strategy	One-leader multi-follower	[215]
Cross-layer anti-jamming design	Discrete strategy	One-leader one-follower	[224]
Deception-based defense	Continuous strategy	One-leader one-follower	[228], [258]

a deception-based defense strategy was investigated to fight against the jamming attack.

In [191], [217] and [253], they extend to multi-user scenarios, and the one-leader multi-followers Stackelberg anti-jamming power control game was formulated. In [210], based on the Stackelberg game framework, the power control problem was analyzed in multi-jammer scenarios. In [211], [212] and [200], the power control problem was investigated in cooperative wireless networks, and the three-layer Stackelberg game framework was adopted, in which relay nodes act as vice leader, and it forms leader-vice leader-follower Stackelberg anti-jamming game framework. In [218], the three-layer secure offloading Stackelberg game was investigated to cope with jamming attacks.

Discrete problems also have some important applications, in which convex optimization techniques are intractable. For these scenarios, intelligent learning method can be employed to obtain the desirable strategies by trial-and-error with jamming environment, and it has attained extensive attention. In [213], [214]and [220], based on Q-learning, the desirable power strategy can be achieved. In [221], based on deep neural networks, the optimal power allocation strategy was obtained for both cluster head node and jammer in sensor edge cloud. In [254], the unmanned aerial vehicle (UAV)-aided anti-jamming problem was investigated in maritime communication, and optimal anti-jamming power strategies were obtained for UAV and transmitting ship. In [255], a DRL-assisted power control anti-jamming scheme was proposed with one smart jammer and multiple eavesdroppers. Based on stochastic learning mechanism, the anti-jamming channel selection problem was discussed in [183] and [245], and mutual interference and malicious jamming were simultaneously considered. In [219], a learning-based dynamic spectrum access scheme was designed, and the log-linear learning algorithm was employed. In [222], a multi-leader one-follower anti-jamming channel selection game was formulated, and an active attraction based learning algorithm was developed to cope with tracking jammer. In [223], the joint channel selection and power control problem was considered in multi-user scenarios, and a multi-leader one-follower Stackelberg anti-jamming game was formulated. In [225], the formulation of anti-jamming subnetwork formulation was analyzed in satellite–enabled army IoT, and the RL-based anti-jamming algorithm was designed. In [226], [227], anti-jamming routing selection schemes were proposed, and DRL was employed to achieve desirable routing paths. In [215], deep Q-networks were employed, and the optimal trajectory was obtained with a UAV jammer. In [224], based on hierarchical learning method, the cross-layer anti-jamming mechanism was investigated, and the routing, channel selection and power control were jointly analyzed.

3)**Stochastic anti-jamming game**

Stochastic game, also known as Markov game, can be regarded as a generalization of the Markov decision process (MDP) with multi-agent cases, and it is a mathematical framework for multi-agent decision optimization in dynamic environment. Mathematically, a stochastic anti-jamming game can be denoted as $\mathcal{F} = \{\mathcal{N}, \mathcal{S}, \mathcal{A}_1, ..., \mathcal{A}_N, r_1, ..., r_N, q\}$, where $\mathcal{N} = \{1, .., N\}$ is the user set, $\mathcal{S}$ denotes the states set, $\mathcal{A}_n, n = \{1, .., N\}$ and $r_n, n = \{1, .., N\}$ respectively represent the strategies set and reward of user n ,and q is the state transition model. In anti-jamming field, it has many advantages. First, it can describe dynamics due to dynamic jamming strategies, such as jamming channel and jamming power. Second, it can characterize the collaborative and competitive relationships among players. Based on the characteristics of utility functions, stochastic game can be divided into three Modes: completely collaborative, completely competitive, and mixed mode [260]. If all agents have the same utility functions, and the stochastic game is a completely collaborative mode. For two agents scenarios, if they have opposite utility function, and it can be regarded as completely competitive mode. If utility functions of agents are neither identical nor opposite, it is considered as a mixed mode. A framework of stochastic anti-jamming game is shown in Fig. 4. A large variety of applications can be found for stochastic anti-jamming game, and they are reviewed in Table VI.

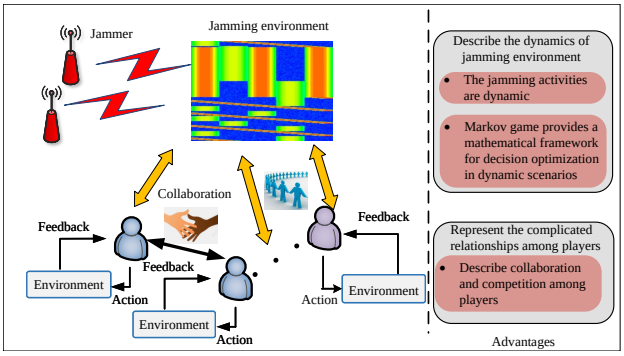

Fig. 4. A framework of stochastic anti-jamming game.

Collaborative mode is common in anti-jamming field, and it has some applications. In [261]–[264], based on stochastic game framework, multi-agent RL anti-jamming methods were proposed, and standard Q-learning was employed. In [185]–[187], collaboration mechanism (e.g., information exchange) was considered to realize collaborative learning among users, and collaborative multi-agent RL algorithms were proposed to simultaneously cope with mutual interference and malicious jamming. In [188], a novel cross check Q-learning method was proposed, and each agent can able to predict the behaviors of other agents. For this scenario, collaborative learning can be realized by behavior prediction. To deal with the limitation of dimensions of state space in large-scale networks, the stochastic game and DRL can be combined. In [229], the anti-jamming channel selection problem was investigated in self-organizing networks, and a decentralized DRL anti-jamming scheme was proposed. In [230], based on deep neutral networks, a mean field RL anti-jamming method was proposed for ultra-dense networks. In [265], a win or learn fast Q-learning anti-jamming method was designed to battle sweep jamming in control channels. In [266], a multi-agent layered Q-learning approach was designed for UAV anti-jamming communication networks, and a two layers framework to respectively find the optimal channel and power strategies.

The completely competitive mode can well describe the

TABLE VI
EXISTING APPLICATIONS OF STOCHASTIC ANTI-JAMMING GAME

Application	Mode	Ref.
Channel selection	Collaborative mode	[185]–[187], [229], [230], [261]–[265]
Joing channel selection and power control		[266]
Channel selection	Competitive mode	[231], [267]–[269]
Power control, channel selection		[232], [233]
Joint adaptation of frequency hopping and transmission rate		[236], [237]
Power allocation		[238]
Path selection		[270]
Joint adaptation of frequency and operation mode		[271]
Channel selection	Mixed mode	[234], [235], [272]
Wireless energy transfer		[273]

confrontation relationship between legitimate user and jammer. In [231], a stochastic anti-jamming zero-sum game in CR networks was formulated, and an anti-jamming defense scheme based on minimax-Q learning was proposed. In [267], a stochastic zero-sum game was employed to capture the competitive interactions between secondary user and jammer, and stationary saddle-point strategies were provided. In [268], the smart hopping was investigated, and inference and logical reasoning were adopted to predict the rival's strategies. In [269], the anti-jamming spectrum auction scheme was designed in CR networks, which can be converted into two-level auctions. In [232], [233], the anti-jamming problems in energy harvesting networks and CR networks were discussed, and two multi-agent RL algorithm (i.e., minimax-PDS and WoLF-PDS) were proposed to quickly learn in dynamic jamming environment. In [236], [237], the joint frequency hopping and transmission rate anti-jamming scheme was discussed, and the constrained NE was analyzed. In [238], a finite-energy anti-jamming problem was considered, and a dynamic programming algorithm was proposed to find the NE solutions. In [270], defensive path selection problem was analyzed, and the optimal path selection mechanism was designed. In [271], the joint adaptation of frequency hopping and operation mode was analyzed in CR networks, and optimal defense strategies were developed to fight against a reactive sweep jammer.

In addition, the mixed mode also has some applications. In [272], a general-sum stochastic game was formulated to tackle control channel jamming problem, and win-or-learn-fast principle based learning scheme was designed to obtain the optimal control channel allocation strategies. In [234], a stochastic anti-jamming game was employed to model interactions between a secondary user and intelligent jammer in CR networks, and optimal frequency hopping schemes were derived. In [235], based on Q-learning, a game-theoretic frequency hopping scheme was designed to select available channel in wireless sensor network. In [273], a constrained stochastic jamming game was analyzed in wireless powered communication networks, and the best response dynamics based iterative algorithm was given to obtain the stationary policies.

4) **Zero-sum anti-jamming game**

The zero-sum game framework involves two players, which have completely opposite utility functions. One player aims to maximize its utility function, while another player is to minimize the same utility function. That is to say, the payoff of one player leads to the loss of the other player in the formulated game-theoretic framework. It can well capture the adversarial relationship between legitimate user and jammer, and has attained extensive attention in anti-jamming field. An anti-jamming zero-sum game contains two players: a legitimate user and a jammer. On the one hand, the legitimate user is to maximize a pre-specified utility function with its own optimal strategy. On the other hand, the jammer aims to minimize the pre-specified utility function from its perspective.

The zero-sum anti-jamming game framework has many applications in the existing literature. First, the power domain anti-jamming defense was investigated in [239]–[241], [274]. In [239], the power control anti-jamming problem was investigated between the legitimate user and jammer, and the capacity was defined as the utility function. The authorized user aims to maximize its capacity under hostile jamming, while the jammer tries to minimize the capacity of the legitimate user. In [241], the power allocation problem between a CR transmitter and a jammer was formulated as a power allocation game, and an optimal power strategy based on Q-learning was obtained for the smart jammer scenario. In [274], the power allocation anti-jamming game was designed in a training-based MIMO system, and the legitimate user jammer and have opposite objectives for data rate. In [240], the frequency hopping jamming game was formulated in a satellite communication network, and optimal power strategies were shown for three scenarios: complete information game, jammer-biased game and defender-biased game.

Second, frequency domain anti-jamming defense scheme can be analyzed by zero-sum game framework. In [242], the frequency domain anti-jamming mechanism was analyzed, and a frequency hopping strategy was considered to cope with the

jamming attacks. A measurement-driven anti-jamming game framework was formulated between the legitimate link and the jammer, in which the legitimate link pursued to maximize its throughout, while the jammer was to minimize this throughout. In [275], a channel selection anti-jamming method was investigated in CR network, and a zero-sum anti-jamming game framework was modeled between the secondary user and jammer. In [276], a hopping pattern selection problem was investigated, and transmission rate was regarded to as a metric to decide on random hopping or spreading. In [231], [267] and [269], based on the zero-sum game framework, the channel selection anti-jamming problem was discussed in CR networks.

Besides, there exist other applications of zero-sum game framework. In [236], [237], the multi-domain anti-jamming of frequency hopping and transmission rate was investigated. In [277], as a counter-jamming approach, energy harvesting was investigated to mitigate jamming attacks. In [278], security defense was considered in underwater wireless networks, and an energy-depleting jamming game was formulated. In [279], a jamming game between the encoder-decoder and jammer was modeled, and the encoder-decoder attempt to choose optimal encoding and decoding strategies to minimize the probability of error. In [280], the adaptation mechanisms for a transmitter/receiver with multiple parameters (e.g., transmission rate, power) were analyzed in a power constrained jammer scenario. In [281], a geometry-based anti-jamming theoretical framework was formulated in underwater sensor networks, and multi-dimensional anti-jamming strategies (e.g., modulation and coding scheme, power strategy) were given for a blind and a reactive jammer. In [270], the anti-jamming mechanisms at network layer were analyzed.

5) **Graphical/Hypergraphical anti-jamming game**

The graphical game [282], as an appropriate mathematical tool, can model the mutual interference effect among legitimate users. It can well describe the spatial distribution characteristics, and a mutual strong interference relationship exists between two neighboring users. In some dense wireless network scenarios, not only strong interference relationships need to be considered, but also the accumulative weak interference relationship cannot be ignored among three or more legitimate users, which can equal to a strong interference relationship when it exceeds a threshold [184]. As an extension of traditional graphical model, hypergraph can accurately model the interference relationships, and both strong interference relationships and cumulative weak interference relationships can be captured [184], [283].

A framework of a hypergraph-based anti-jamming game is shown in Fig. 5. The strong interference relationships are denoted as lines, and the accumulative weak interference relationships are expressed as circles. For example, there is a strong interference relationship between user 1 and user 2 if the same channel is adopted. Similarly, strong interference relations are formed between user 2 and user 5, and between user 2 and user 6. Besides, there is a cumulative weak interference relation among user 1, user 3, and user 6, and among user 4, user 5, and user 6. Specifically, user 1 and user 3 do not interfere with user 6 alone, while user 1 and user 3 together interfere with user 6 if the same channel is employed. The hypergraph-based anti-jamming game can fully capture the interference relationships, and strong interference relations and weak interference relations can be simultaneously described. In addition, the formulated hypergraph game can analyze and model the anti-jamming problem, which can be transformed into a generalized interference and jamming minimization problem to obtain the optimal anti-jamming strategies. A graph-based anti-jamming game can be considered as a simplification of hypergraph-based anti-jamming game, and only strong interference relations are considered.

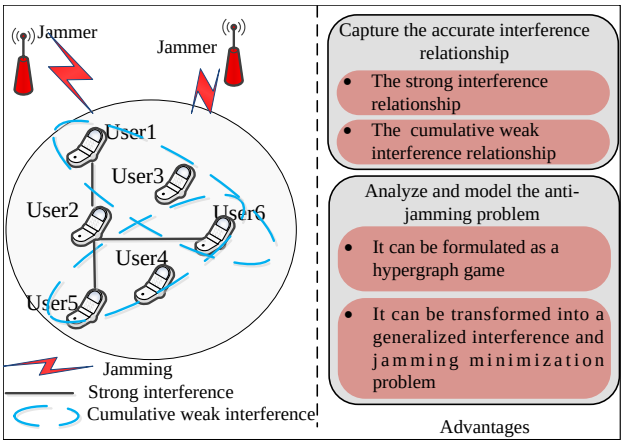

Fig. 5. A framework of hypergraph-based anti-jamming game.

The graph-based anti-jamming game is a promising method in dense wireless networks, and it has attracted some attention. In [243], [244], graph model was considered to characterize the interference relations between two neighbor users, and a context-aware anti-jamming cannel access mechanism was designed. Then, a distributed learning algorithm was designed to obtain NE solutions for different jamming patterns. In [184], a hypergraph-based anti-jamming spectrum access scheme was proposed in dense wireless networks, and hypergraph model was employed to accurately describe the interference relations among legitimate users.

6) **Other anti-jamming games**

1) **Colonel Blotto anti-jamming game**

Colonel Blotto game is an extension of the non-cooperative game, and it involves two opposing players, which allocate limited resources on N independent battlefields [284]. The player 1 (player 2) with more powerful strategies can win the battlefield, and the utility function of each player is the sum outcomes from all battlefields. For a Colonel Blotto game, each player allocates its forces without other player's information, and it has some studies in anti-jamming field.

In [285], [286], the power allocation problem was analyzed between the secondary user and jammer in CR networks. In [287], a jamming power game was formulated between a controller node and a jammer for OFDM-based IoT networks. In [288], an anti-jamming Colonel Blotto game was formulated to model the confrontation problem between a fusion center and jammer in IoT. In [289]–[292], the anti-jamming power control strategy was designed for a health monitoring system while considering the limited power constraints as well as the multichannel fading. In [293], a power control anti-jamming

scheme was investigated in IoT networks, and a heterogeneous iterative method was proposed to obtain the desirable solutions. In [294], a Colonel Blotto game was formulated between a secondary access point and a jammer in a dynamic spectrum access network. In [295], [296], the attack-defense problem was modeled as the networked Colonel Blotto game in the network system, and players (e.g., attacker and defender) aimed to maximize or minimize network performance.

2) **Anti-jamming relay game**

Relay-assisted anti-jamming game is a powerful tool to cope with jamming attacks in wireless communications, and the relay can help information transmission in jamming networks. It improves the anti-jamming performance from the perspective of relay enhancement, and has some applications in the existing literature.

In [200], [211]and [212], the power control anti-jamming problem was investigated in cooperative wireless communication networks, and relay node was employed to help information transmission. In [297], [298], an anti-jamming relay game was designed in UAV-assisted vehicular ad-hoc networks, and UAV was regarded as a relay. The UAV acts as a defender to decide whether or not to relay the information transmission from an onboard unit to another jamming-free roadside unit when the serving roadside unit suffered from heavy jamming, while the jammer chooses its optimal power strategy. In [299], a power control game was formulated between the relay UAVs and the jammer in a UAV network, and multiple UAVs act as relays to help the transmission of the resource-destination link. In [345], an anti-jamming scheme was designed for UAV-aided cellular systems, and the UAV acts as a relay when the base station suffers from heavy jamming. The UAV can choose its optimal power strategy to fight against jamming attacks, and the smart jammer can optimize its jamming power strategy to minimize the utility of the UAV.

3) **Prospect-theory based anti-jamming game**

In traditional anti-jamming game frameworks, all players are assumed to be rational, and immune to real-life perception [300], [301]. In these scenarios, the legitimate users and jammers make their decisions based on their expected utilities under uncertainty. However, this assumption cannot characterize the subjectivity of players. Fortunately, prospect theory (PT) has emerged as an appropriate tool to analyze anti-jamming game from a user-centric view, and a probability weighting function is adopted to describe the subjective decision process of players.

In [300], a PT-based anti-jamming game was formulated between a secondary user and a jammer, and the channel access problem was investigated in CR networks. In the formulated game, each player has a subjective view on random action of another player, and each player selects its channel strategy to maximize its PT-based throughout. Then, the authors extended to PT-based power control game scenarios under the uncertainty of channel gains and the strategy of opponent in [301], and subjective secondary user and jammer select their power strategies to maximize their SINR. In [302], a PT-based smart attack game was analyzed between a UAV and a smart attacker, and the subjectivity of smart attack was described under the uncertainty of detection accuracy. In [303], a PT-based cloud storage defense game was investigated between an attacker and a store defender.

4) **Evolutionary anti-jamming game**

Evolutionary game is an appropriate mathematical framework to analyze the interaction behaviors among agents in a population [40], and it is an extension of non-cooperative game by introducing the concept of populations. Replicator dynamics is employed to analyze the evolutionary stable strategy (ESS) strategies, and can model the evolutionary process of population over time. To be specific, mutation and selection mechanisms are adopted to realize self-replication and elimination. It is suitable for analyzing collective and irrational behaviors of agents, and can describe the dynamic evolution process between the legitimate users and jammers.

In [304], an evolutionary anti-jamming game framework was employed to analyze anti-jamming problem in a cooperative network, in which there were M users and N jammers. The M users aim to collectively maximize their SINR, while N jammers try to degrade the SINR of users. Their strategies are to either transmit or not with some probability. Then, based on replicator dynamics, the evolution of ESS strategies was presented for different cooperation levels of populations. In [305], an evolutionary game based anti-jamming channel selection scheme was designed in CR networks. In [306], an evolutionary power control game was investigated to counteract responsive and non-responsive jamming attacks. In [307], the anti-jamming problem was investigated in NOMA system, and an evolutionary anti-jamming game was established between the base station and jammer. Then, learning based algorithms were designed to obtain the desirable power strategies in dynamic jamming environment.

Remark 2. *The above game models have their own unique advantages from different perspectives, and can well analyze and model the anti-jamming problem in some anti-jamming scenarios. In addition, other game models can also be found in some anti-jamming scenarios. In [308], [309], differential game framework was adopted to analyze anti-jamming communication in an UAV network, and optimal-control theory was developed to achieve the optimal equilibrium solutions in dynamic jamming environment. In [310], psychological behavior was captured in Internet of Battlefield Things, and an anti-jamming scheme was designed from the perspective of psychological game. In [311], a bimatrix game framework was applied to analyze the anti-jamming problem in frequency hopping communications.*

C. Strengths and limitations of game theory-based anti-jamming approaches

Based on the above analyses of game theory in anti-jamming problem, its strengths can be given as follows:

(1) Owing to the inherent confrontational feature, there exist adversarial jamming relationships between the legitimate users and malicious jammers. Moreover, there are also competitive mutual interference relationships among legitimate users in multi-user scenarios. Game models can well capture and analyze jamming relationships between the legitimate users

and jammers, and mutual interference relationships among legitimate users.

(2) Combined with intelligent learning technologies, the game theoretic learning framework can obtain desirable anti-jamming strategies, especially in dynamic, incomplete and unknown information constraints. The game theoretic learning framework will be discussed in Section V.

However, game theory has some limitations in anti-jamming field, and it can be summarized as follows:

(1) Game theory based anti-jamming methods usually require some prior knowledge of opponents, such as strategies space, and utility function. Therefore, the legitimate users need to estimate some information of jamming environment, such as environment parameters and jamming pattern. Unfortunately, due to the adversarial relationship between the legitimate users and jammers, it is difficult to obtain the prior information of opponents, and information loss will be inevitable in dynamic, incomplete and unknown jamming environment due to the estimation of jammer information.

(2) For some advanced jamming patterns with intelligent characteristics, dynamic and intelligent jamming can be created. As a consequence, it is difficult to estimate and model the jamming strategies and utility function from the perspective of engineering realization. Besides, it is difficult to make real-time response when the jammer changes its jamming strategies.

IV. Reinforcement learning based anti-jamming methods

As an important sub-field of machine learning, RL has the advantage of learning in unknown environments, and it can learn by interactions with the environment. It is rooted in the study of animal behavior, and can realize the mapping from environment state to action. The desirable strategies can be found by trial and error interactions, and better actions can obtain higher reward, while bad actions can be punished with lower reward. In anti-jamming field, RL can deal with the dynamic and unknown information constraints due to adversarial relationship in jamming environment.

In this section, the basic model of RL is firstly introduced, and the some anti-jamming RL applications are analyzed and reviewed. Finally, strengths and limitations are presented.

A. Basic model of reinforcement learning

For a basic model of RL, it can be cast as a Markov decision process (MDP), and includes four elements: state, action, reward and state transition probability function, and it can be expressed as a four-element tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}\}$, where $\mathcal{S}$ represents the state set, $\mathcal{A}$ is the action set, $\mathcal{R}$ denotes the reward, and $\mathcal{P}$ is the state transition probability function.

- **State**: Each state $s \in \mathcal{S}$ represents the description of perceived environment. At each time step of interactions, the system is in some state $s \in \mathcal{S}$.
- **Action**: It can be regarded as a mapping from state to action. Based on the current state $s \in \mathcal{S}$, the agent chooses an action $a \in \mathcal{A}$.
- **State transition probability function**: It is a probability distribution function over state set $\mathcal{S}$. If an action $a \in \mathcal{A}$ executes on the current state $s \in \mathcal{S}$, it makes the system move into a new state $s' \in \mathcal{S}$. The state transition function is usually unknown in jamming environment, and RL can be used to solve the MDP problem, such as Q-learning.
- **Reward**: Based on the current state and action, the learning agent can obtain a corresponding reward from the system. As time evolves, the good actions should have higher reward, and the bad actions should be punished with a lower reward. Appropriate reward function makes the learning process tend to the desirable direction, such as maximum throughput, and minimal jamming level.

B. The applications of RL in anti-jamming communications

RL has shown significant strength to solve anti-jamming problem, and various applications can be found in existing work, such as Q-learning, multi-armed bandit (MAB), DRL and transfer RL.

1) Q-Learning

As a typical RL method, Q-learning is an effective tool [312], and it has attained widespread attention in the anti-jamming field, such as channel selection and power control. The state-action table is called Q table, and the agent takes actions based on it. At every time step, the Q table can be updated according to the received reward, which can be obtained from the interactions with the dynamic and unknown jamming environment. Different from game theory based anti-jamming solutions, jamming behaviors can be learned by interactions with environment without estimating jamming parameters and patterns. The existing studies are reviewed and compared in anti-jamming field in Table VII.

First, Q-learning has some applications for channel selection anti-jamming mechanisms. In [52], a Q-learning based anti-jamming method was designed to pro-actively avoid jamming channel under the condition of one sweeping jammer. In [53], a cooperative Q-learning method was proposed to avoid jammed channel, and meanwhile it can solve the hidden jammer problem that may actually jam information transmission, but it was not detected by the learning node. In [313], a Q-learning based dynamic spectrum anti-jamming algorithm was given to achieve the desirable channel selection strategy in fading environment. In [314], Q-learning based jamming avoidance scheme was presented for wideband autonomous CR networks with one sweeping jammer scenario. In [315], a collaborative UFH-based broadcast method was proposed, and an optimal frequency hopping strategy was achieved. In [261]–[264], [185]–[188], they extended channel selection anti-jamming problem to the multi-agent scenarios. In [261]–[264], multi-agent RL anti-jamming schemes were investigated, and each user aimed to evade the interference of other users as well as avoid malicious jamming. In [185]–[187], collaboration was considered by information exchange, and collaborative multi-

TABLE VII
SUMMARY OF RL MODELS IN ANTI-JAMMING FIELD

Application	Key contribution	Method	Ref.
Channel selection	Avoid jamming channel of sweeping jammer.	Q-learning	[52], [53] [313], [314]
	Achieve optimal frequency hopping strategy.	Q-learning	[315]
	Evade the mutual interference as well as avoid jamming.	Multi-agent RL	[261]–[264]
	Cope with the mutual interference and malicious jamming simultaneously.	Collaborative multi-agent RL	[185]–[188]
Power control	Obtain the optimal power strategy with unknown network parameter scenarios.	Q-learning	[212], [316]
	Obtain optimal power strategy to deal with sweeping jammers and smart jammers.	Q-learning	[317]
	Achieve the optimal power strategy without knowing the jamming model and channel model.	Policy hill-climbing algorithm	[318]
	Add memory component into classic Q-learning to maximize the total capacity.	Q-learning	[319]
Time-domain anti-jamming	Enable the legitimate user to switch between "active" and "silent" state to escape random pulse jamming attacks.	Q-learning	[119]
Ambient backscatter	Not only can escape the jamming attacks but also can leverage jamming signals.	Q-learning	[320]
Multi-dimensional anti-jamming	Jointly exploit the multi-dimensional "frequency-motion-antenna" space to improve the anti-jamming performance in UAV swarms.	Q-learning	[321]
Joint optimization of power allocation and reflecting beamforming.	Jointly optimize the power allocation and reflecting beamforming in IRS-assisted anti-jamming communication.	WoLF-CPHC	[322], [323]
Joint optimization of channel selection and data scheduling	Jointly optimize the channel selection and data scheduling in HF jamming environment.	Q-learning	[324]
Jamming deception	Fool the jamming by deceiving it into attacking a victim channel to secure communication in other safe channels.	Successive RL	[155]

agent RL anti-jamming mechanisms were proposed to simultaneously cope with mutual interference among legitimate users and malicious jamming. Instead of information exchange, a distributed multi-agent RL was designed to address jamming and avoid interference in [188], and cooperation was realized by an estimation of Q-table of other legitimate users.

Second, Q-learning was also widely adopted in the power control anti-jamming problem. In [212], the RL based anti-jamming algorithm was provided to obtain the optimal power strategy for cooperative CR networks with unknown network parameters scenarios (i.e, channel gains and transmission cost). In [316], a Q-learning based power control anti-jamming scheme was proposed to address jamming attacks for underwater sensor networks with unknown channel parameters. In [317], based on Q-learning, a power control anti-jamming method was investigated in heterogeneous CR networks. In [318], a fast policy hill-climbing algorithm, a modified Q-learning method, was employed to achieve optimal power strategy for mmWave massive MIMO system. In [319], memory component was added into classical Q-learning, and a modified Q-learning was proposed to maximize the total capacity with a smart jammer scenario.

Besides, Q-learning has many other anti-jamming applications. In [119], a time-domain anti-jamming method was proposed to fight against random pulse jamming, and the desirable strategies can enable the legitimate users to switch between "active" and "silent" state to escape jamming attacks. In [320], ambient backscatter technology was adopted, and the transmitter not only can escape the jamming attacks but also can leverage jamming signals. Then, a Q-learning based algorithm was presented to achieve the optimal operations. In [321], the multi-dimensional degree-of-freedom in "frequency-motion-antenna" space was jointly optimized, and a Q-learning based anti-jamming approach was proposed to improve the anti-jamming performance in UAV swarms. In [322], [323], the anti-jamming problem was investigated in intelligent reflecting surface (IRS) assisted communication, and a fuzzy win or learn fast-policy hill-climbing (WoLF-CPHC) learning method was designed to solve the joint optimization problem of power allocation and reflecting beamforming strategy. In [324], the joint channel selection and data scheduling problem was investigated for high-frequency (HF) communication in jamming environment, and a modified RL anti-jamming mechanism was proposed, in which it combined classical Q-learning with upper confidence bounds to deal with the large action set space. In [155], the jamming deception problem was investigated, and the jammer was deceived into attacking a victim channel to secure communication in other safe channels. Then, successive RL-based algorithm was presented to find the optimal power and channel strategy.

2) **Multi-armed Bandit**

As a kind of stateless RL, MAB can solve the decision optimization problem in dynamic and unknown environment, and it can model interactions between learner and environment [325]. A MAB-based learning framework consists of an action space and reward function, and the agent can deal with a dilemma of K actions. Each action is regarded as an arm, and different actions yield different rewards. The regret learning is the basic learning framework, and the performance metric usually employs the regret R(t), which is the performance difference between the received actual reward and the expected optimal reward. The regret represents the reward loss, and it is because the system does not always adopt the optimal action. The objective of strategy optimization is to minimize the term "regret". Correspondingly, the optimization problem can be formulated as a regret minimization problem. Generally, the MAB models can be categorized into three major types, i.e., stochastic, adversarial and combinatorial [141], [182]. In stochastic MAB models, the reward is generated with an independent and identically distributed (IID) process. While, in adversarial MAB models, the reward is arbitrary, and its generation distribution is not IID. For combinatorial MAB model [182], it bridged the stochastic and non-stochastic MAB problems into a unified combinatorial MAB framework, and it no longer restricted the distribution model. The MAB models have gained growing attention in anti-jamming field, and the existing studies are reviewed and compared in Table VIII.

The stochastic MAB model is the basic case, and the received reward follows an IID process with a fixed unknown distribution. In [142], a multi-domain anti-jamming framework was formulated in the power and spectrum domain, and a UCB1 based anti-jamming scheme was proposed with unknown channel state information. In [326], [327], a MAB analytical framework was developed for competing CR networks, and it jointly coordinated own communication activities and jamming their opponents. Then, the spectrum access problem was modeled as a stochastic MAB framework, and a Thompson sampling based optimal spectrum access algorithm was provided with unknown channel parameters.

It should be pointed out that the reward depends on the IID assumption for the stochastic MAB model, and this assumption does not hold in some jamming scenarios. However, the adversarial MAB model has no restrictions on reward, and it has attracted increasing attention in practical anti-jamming applications. In [328]–[330], the anti-jamming channel access problem was formulated as an adversarial MAB framework in CR networks, and online jamming-resistant channel access algorithms were developed with unknown channel statistics. In [140], a multi-domain anti-jamming problem was formulated as an adversarial MAB model for aeronautic swarm tactical network, and a KL-UCB^{++} based algorithm was provided to obtain configuration strategy in power and spectrum domain with dynamic and unknown aeronautical swarm network environment. In [141], the jamming defense problem was formulated as an adversarial MAB framework for remote state estimation in cyber-physical systems, and an online-learning based anti-jamming scheme was designed to jointly choose the optimal channel and power strategy without prior knowledge of channel state information and the jamming strategy. In [331], the utility optimal scheduling problem was investigated in multi-hop wireless networks, and a cross-layer anti-jamming mechanism was proposed to counteract the reactive jamming. Then, based on EXP4 algorithm, a jamming-aware online learning algorithm was presented to obtain the utility optimal cross-layer solution in physical, link and routing layers, and it is robust to varying jamming behaviors. In [332], the UFH-

TABLE VIII
SUMMARY OF MAB MODELS IN ANTI-JAMMING FIELD

Application	Key contribution	Method	Models	Ref.
Muti-domain anti-jamming	A multi-domain anti-jamming framework was formulated, and a MAB based algorithm was proposed with unknown jamming environment.	UCB1	Stochastic	[142]
Channel selection	The optimal spectrum access algorithm was provided with unknown channel parameters for competing CR networks.	Thompson sampling		[326], [327]
Channel selection	Anti-jamming channel access algorithms were presented with unknown channel statistics.	EXP3 algorithm	Adversarial	[328]–[330]
Muti-domain anti-jamming	Multi-domain anti-jamming strategy in power and spectrum domain can be obtained with dynamic and unknown aeronautical swarm network environment.	KL-UCB^{++}		[140]
Muti-domain anti-jamming	Jointly choose the optimal channel and power strategy without prior knowledge of channel state information and the jamming strategy.	EXP3 algorithm		[141]
Utility optimal scheduling	A jamming-aware online learning algorithm was presented to obtain the utility optimal cross-layer solution in physical, link and routing layers.	EXP4 algorithm		[331]
UFH	An online learning based UFH algorithm was designed to address oblivious and adaptive jammers.	EXP3 algorithm		[332]
Channel selection	Find the optimal channel strategy, and achieve near-optimal learning performance without any prior knowledge of environment.	EXP3 algorithm	combinatorial	[182]
Spectrum aggregation and access	Find the near-optimal solution without prior knowledge of channels and jammers.	EXP3 algorithm		[333]
Shortest path routing	Achieve the near-optimal performance without any prior system knowledge.	EXP3 algorithm		[334]

based anti-jamming problem was formulated as an adversarial MAB framework, and an online adaptive learning algorithm was designed to address the oblivious and adaptive jammers.

Generally speaking, the stochastic and adversarial MAB frameworks rely on distinctively different analytic methods and have different performance. Fortunately, the combinatorial MAB framework has a unified framework and can be applicable for both stochastic and adversarial regime, and it is highly desirable in practical anti-jamming communication applications. In [182], the authors proposed combinatorial MAB framework for the first time, and an EXP3 based algorithm was presented to find the near-optimal channel access strategy without any prior knowledge of jamming environment. The designed algorithm can obtain near optimal performance for both stochastic and adversarial regime. In [333], the jamming-resistant spectrum aggregation and access problem was investigated, and an online learning algorithm was proposed to obtain an effective solution in CR network. In [334], the shortest path routing problem was formulated as a combinatorial MAB framework in jamming environment, and an innovative EXP3 based algorithm was provided to achieve near-optimal performance without any prior system knowledge.

3) **Deep reinforcement learning**

Q-learning has been widely applied in anti-jamming field, and it has obtained widespread attention. However, the size of state space is closely related to the performance of Q-learning, and its weakness is exposed when the state space is large. The larger the state space, the slower the converge speed. Therefore, due to the curse of dimensionality, the high-dimensional state spaces pose great challenges to typical Q-learning method. Fortunately, DRL is an enabling technology that can improve the learning ability [335], and can effectively address the curse of dimensionality. It is an integrated framework of the RL and deep learning, and deep neural networks are employed to extract useful features and approximate Q-function. The neural networks are adopted to approximate and replace Q-function, such as convolutional neural network (CNN), fully connected network (FCN), and recurrent neural network (RNN). The DRL empowered anti-jamming schemes will obtain desirable performance with large state space in dynamic and unknown jamming environment. The DRL has drawn extensive investigations in anti-jamming field, and the existing anti-jamming schemes are reviewed and compared in Table IX.

First, DRL-based spectrum domain anti-jamming schemes have attracted significant attention, and various kinds of countermeasures were proposed. In [180], based on a recursive CNN, a DRL based anti-jamming mechanism was proposed in a dynamic and unknown jamming environment, and the spectrum waterfall was directly regarded as a state. In [181], to address a large number of action spaces in broadband networks, a hierarchical DRL based anti-jamming approach was presented with unknown jamming patterns and channel model, and a two-level action selection framework was established. Specifically, the frequency band is chosen at first, and then the specific frequency is chosen from the selected frequency band. In [336], a pattern-aware DRL anti-jamming approach was proposed to obtain the optimal channel strategy with changeable jamming patterns, and a sliding window mechanism and deep learning were employed to identify jamming patterns. In [337], a DRL-based robust anti-jamming spectrum access mechanism was designed with incomplete sensing information, and a generative adversarial network was employed to complete missing spectrum information. In [338], a primary user-friendly anti-jamming spectrum access mechanism was designed in overlay CR network, and both offline traing and online deploy were considered. In [339], a DRL based anti-jamming channel selection method was presented with heterogeneous information fusion in HF communication network, and a composite jamming environment state was considered. Then, a new deep Q-network framework was designed, in which the CNN was employed to process spectrum state, and the FCN was employed to process channel gain state. In [340], a double deep Q-network based anti-jamming mechanism was presented to obtain the optimal channel strategy with a partially observable environment in a heterogeneous wideband spectrum network. In [341], a DRL-based defense strategy was designed to confront a RL-based intelligent jammer. In [342], a Transformer Encoder Q network was formulated, and a double deep Q network based anti-jamming mechanism was presented in CR network to fight against various jamming attacks, such as sweep jamming and random jamming. In [343], based on feature engineering, an improved anti-jamming mechanism was developed, and an improved state space was employed to reduce the computational complexity.

In [173], [229], [230]and [344], the DRL based anti-jamming mechanism was extended to multi-user scenarios, and various multi-user DRL anti-jamming algorithms were proposed. In [344], the double deep Q-learning anti-jamming algorithm was presented for a multi-user system model, and the performance of three networks model, i.e, CNN, FCN and long short term memory (LSTM), were evaluated. In [229], a decentralized DRL based anti-jamming method was presented for self-organizing networks. In [230], a mean field DeepMellow based anti-jamming spectrum access scheme was proposed for ultra-dense IoT networks. In [173], a DRL based anti-jamming scheme with continuous action space was proposed in an ultra-dense network without estimating the jamming environment.

Second, DRL can be applied in power domain anti-jamming measures, and DRL-based power control anti-jamming schemes were investigated. In [345], the DRL based anti-jamming scheme was presented to choose the optimal power strategy in UAV-aided cellular networks, without a prior knowledge of network topology, message generation model, server computation model and jamming model. In [346], a deep Q-network based anti-jamming power control scheme was designed in IoT networks, and the optimal transmission power strategy can be determined with unknown network topology and jamming model. In [347], a deep deterministic policy gradient based anti-jamming scheme was formulated in UAV networks, and the optimal power strategy can be acquired with no prior information of jamming model and jamming power.

Third, DRL can be employed in multi-domain ant-jamming

TABLE IX
SUMMARY OF DRL-BASED ANTI-JAMMING APPROACHES

APPlication	Key contribution	Model	Ref.
Channel selection	Obtain the optimal spectrum access strategy without estimating jamming pattern and parameters.	CNN	[180]
	Obtain the optimal spectrum access strategy in broadband communication with unknown jamming pattern and channel model.	CNN	[181]
	Obtain the optimal channel strategy with changeable jamming patterns.	CNN	[336]
	Obtain the optimal spectrum access strategy with incomplete sensing information.	CNN	[337]
	Obtain the optimal spectrum access strategy in overlay CR networks.	CNN	[338]
	Obtain the optimal channel strategy with heterogeneous information fusion in HF communication.	CNN,FCN	[339]
	Obtain the optimal channel with a partially observable environment in a heterogeneous wideband spectrum network.	CNN	[340]
	Obtain the anti-jamming spectrum access strategy with intelligent jammer.	CNN	[341]
	Obtain the optimal policy with unknown jamming pattern and channel model in CR network.	Transformer	[342]
	Obtain the optimal frequency hopping strategy with an improved state space and channel switch cost.	RNN	[343]
	Obtain the optimal channel strategy in a multi-user jamming environment.	CNN,FCN,LSTM	[344]
	Obtain the optimal spectrum access strategy with a decentralized framework in self-organizing networks.	CNN	[229]
	Obtain the optimal spectrum access strategy for ultra-dense IoT network.	CNN	[230]
Power control	Choose optimal actions with continuous action space in ultra-dense networks without estimating the jamming parameters and patterns.	CNN	[173]
	Choose optimal power strategy with unknown network topology, message generation model, server computation model and jamming model.	CNN	[345]
	Determine the optimal transmission power with unknown network topology and jamming model in IoT networks.	CNN	[346]
	Acquire the optimal power strategy with unknown jamming model and jamming power in UAV networks.	CNN	[347]
Joint channel selection and power control	Choose the power and channel adaptively with no prior information of jamming patterns, moving trajectory and detection threshold.	CNN	[348]
Joint decision of power control and node mobility	Achieve two dimensional joint strategies for underwater acoustic networks with unknown jamming and channel model.	CNN	[54], [349]
Joint decision of channel, power, modulation and coding rate, and video compress encoding rate	Achieve multi-dimensional joint strategies for low-latency video streaming with unknown jamming and channel model.	FCN	[55]
Two-dimensional decision of user mobility and frequency hopping	Obtain the two-dimensional anti-jamming strategy that determines the optimal channel and user mobility with unknown jamming and channel model.	CNN	[350]
Joint design of user-centric clustering, beamforming and artificial noise	Obtain a fast and adaptive response with dynamic and persistent jamming and eavesdropping attacks.	CNN	[351]
Joint rate adaptation and ambient backscatter	Obtain the optimal defense strategy with unknown jamming attacks and ambient RF signals.	FCN	[352]
Joint design of transmission energy and phase shift	Obtain the optimal policy with sensing errors and limited battery capacity.	FCN	[353]
Jamming deception	Obtain the optimal deception strategy with dynamic environment and unknown jammer.	CNN	[151]
	Obtain the optimal deception strategy without knowing jammer information.	FCN	[152], [354]
Routing path	A distributed cooperation network framework was designed to defend against jamming attacks.	FCN	[355]
Trajectory Optimization	Obtain the optimal trajectory to elude UAV jamming attacks with incomplete channel state information.	CNN	[215]

problems, and DRL-based multi-dimensional anti-jamming schemes were developed. In [348], the joint channel and power selection problem was investigated, and a DRL-based hidden strategy was developed to resist intelligent reactive jamming. In [54], a DRL based anti-jamming scheme was provided to obtain the optimal strategy of power control and node mobility in underwater acoustic networks, without knowing jamming model and channel model. In [55], a DRL based anti-jamming low latency video streaming scheme was presented to obtain the optimal multi-dimensional strategy in visual IoT networks, without knowing the jamming and channel model. In [350], based on DRL framework, a two-dimensional anti-jamming scheme was developed to choose the optimal channel and determine whether to leave heavy jamming area with unknown jamming and channel model. In [351], the joint design framework of user-centric clustering, beamforming and artificial noise was investigated in ultra-dense networks, and the communication behaviors were divided into association phase and transmission phase. In the association phase, an optimization based scheme was presented to obtain a fast response from the perspective of short-term. In the transmission phase, a DRL based mechanism was developed to adaptively adjust the beamforming and artificial noise vectors to deal with dynamic and persistent jamming and eavesdropping attacks from the perspective of long-term. In [352], the rate adaptation and ambient backscatter were adopted to defeat jamming attacks, and a deep dueling neural network framework was presented to obtain the optimal strategy with uncertain jamming attacks and ambient RF signals. In [353], the joint design problem of transmission energy and reconfigurable intelligent surface (RIS) phase shifts reconfiguration was investigated in anti-jamming RIS communication, and DRL based method was designed to obtain the optimal policy with sensing errors and limited battery capacity.

Besides, DRL can be exploited to solve other anti-jamming problems, such as jamming deception, anti-jamming routing path and anti-jamming trajectory optimization. In [151], the deception mechanism was investigated to defeat reactive jamming attacks in low-power IoT networks, and the device can send fake signals to lure jamming attacks. Then, a two-module DRL anti-jamming framework was formulated to obtain the optimal deception strategy with dynamic environment and unknown jammer. To further improve deception based anti-jamming performance, a novel DRL anti-jamming algorithm, based on a deep dueling neural network architecture, was proposed to find the optimal deception strategy in [152], [354]. In [355], a jamming-aware routing path scheme was designed, and a distributed cooperation framework was proposed to defend against jamming attacks. In [215] based on the DRL framework, the trajectory optimization problem was investigated, and the ground users can obtain the optimal trajectory with incomplete channel state information in order to elude UAV jamming attacks.

Remark 3. *Due to the powerful learning ability of DRL, it is an effective method to deal with high dimensionality of the state space, and has attracted extensive attention. However, knowledge-based RL is another promising tool to cope with high-dimensional problems. Specifically, it employs domain knowledge to construct a virtual environment, and it can pre-train with virtual environment in advance. By embedding knowledge into the RL framework, the state space can be compressed. In [356], a knowledge-based RL anti-jamming scheme was proposed to address smart jamming attacks for UAV networks, and optimal flight control and power allocation method was presented for target reconnaissance mission.*

4) **Transfer reinforcement learning**

Although RL has been successfully adopted in many anti-jamming applications, it has some limitations in practical anti-jamming scenarios. The ideal RL scenario corresponds to a steady environment and abundant training. The approach needs to learn again from zero experience when the environment is changed into a similar but different new environment. Fortunately, transfer learning has emerged as a desirable learning framework, which aims to reuse learned knowledge from a previous learning task to another new learning task with faster and better solutions [357], [358]. The key motivation of transfer learning in RL is to accelerate the convergence process and reduce the number of training samples needed for a new task through knowledge transfer across tasks. The transfer RL has some anti-jamming applications, and the existing schemes are reviewed and compared in Table X.

Based on Q-learning and transfer learning, the hotbooting based transfer RL anti-jamming schemes were proposed to reduce the convergence time in [298], [359] and [360]. The Q-table can be regarded as transferred knowledge, and the Q-table can be initialized with the learned experience in similar anti-jamming scenarios. Therefore, it can effectively avoid the initialization with an all-zero matrix, and accelerate the learning process. In [298], a hotbooting based fast anti-jamming mechanism was designed in vehicular ad hoc networks, and the optimal relay strategy was obtained. In [359], a fast power control anti-jamming approach was developed in wireless body area networks. In [360], a multi-regional anti-jamming transfer RL scheme was proposed to obtain the optimal channel selection strategy across multiple regions. Moreover, based on the transfer learning and actor-critic RL algorithm, a transfer actor-critic anti-jamming scheme was presented to obtain the optimal channel selection strategy in a CR network in [361], and the learned action and state information knowledge can be transferred from a source task to a target task.

It is another interesting topic to combine DRL and transfer learning in anti-jamming problems, and transfer learning based DRL anti-jamming mechanisms were developed to pursue fast and better solutions in [255], [345], [362] and [363]. The CNN weights were regarded as transferred knowledge, and the CNN weights can be initialized by learned experience in similar anti-jamming scenarios to accelerate the convergence speed and avoid the initial random exploration. In [255], a transfer learning-assisted DRL anti-jamming power control scheme was designed to guarantee reliable transmission with one smart jammer and multiple eavesdroppers. In [345], the hotbooting technology was employed in UAV-aided cellular anti-jamming communication. Then, based on DRL and hotbooting, the optimal power control strategy can be obtained

to fight against jamming attacks. In [364], a safe hierarchical RL approach was proposed for anti-jamming defense, and inter-agent transfer learning was employed for efficient initial learning. In [362], a fast deep Q-network based frequency-spatial two-dimensional anti-jamming scheme was proposed to resist jamming attacks, and the hotbooting technique was employed to utilize the previous experience. In [363], the multi-dimensional anti-jamming decision was investigated in UAV video transmissions communication, and a safe transfer DRL anti-jamming scheme was developed to guarantee the video quality-of-experience and reduce the outage probability.

Remark 4. *Besides the above methods, other learning approaches can be found in the anti-jamming field. In [365], the jamming defense problem was formulated as a pursuit-evasion framework, and no-regret learning algorithm was presented. In [366], the no-regret learning algorithm was employed to analyze the capacity maximization problem in wireless jamming environment. In [367], multi-tasking learning based on an anti-jamming scheme was analyzed, and multiple learning methods were combined to improve the long-term reward in jamming environment. In [368], a Bayesian learning based anti-jamming framework was proposed to tackle cross-layer attacks. In [369], a federated RL based jamming defense method was proposed for flying ad-hoc networks, and the spatial retreat mechanism was designed to choose alternative paths by retreating jammed spaces. However, RL has some unique advantages in anti-jamming field, and we will discuss in Section IV-C.*

C. Strengths and limitations of RL-based anti-jamming approaches

Based on the above analyses of the RL-based anti-jamming schemes, its strengths can be given as follows:

(1) RL-based anti-jamming approaches can learn jammer behaviors in dynamic and unknown jamming environments, and the prior information is not needed for jamming patterns and parameters. Consequently, few assumptions are needed of jammers, and the information loss can be avoided due to jammer estimation.

(2) In RL-based anti-jamming schemes, jammers are acted as environment, and it is not necessary to have accurate utility function of jammers. Some advanced jamming attacks can be modeled by MDP framework.

However, RL-based methods have some limitations in anti-jamming field, and it can be summarized as follows:

(1) They need thousands of iterations to converge to a stable solution. For DRL-based anti-jamming schemes, they need to spend a lot of time training the network.

(2) The RL-based anti-jamming schemes lack theoretical analysis, and the effectiveness is often difficult to guarantee.

V. Open issues and future research

A. The integration of game theory and RL in anti-jamming problems

Although there were lots of studies for game theory based anti-jamming solutions and RL-based solutions in anti-jamming field, the integration of game theory and RL is an open issue. As stated before, game theory provides powerful mathematical tools to analyze and model the interactions in anti-jamming problem, and various anti-jamming game models can be formulated to characterize the adversarial relationships between legitimate users and jammers, and competitive mutual interference relationships among legitimate users. However, game models only provide a theoretical analysis framework. In order to obtain the desirable anti-jamming strategies, other methods are needed, such as convex optimization theory and RL. Fortunately, RL technologies are promising methods to achieve desirable anti-jamming strategies through interactions with jamming environment, and they can deal with incomplete and unknown information constraints. Therefore, the game theoretic learning framework is a natural tool to describe and solve anti-jamming problem. A game theoretic learning framework for anti-jamming communication is shown in Fig. 6.

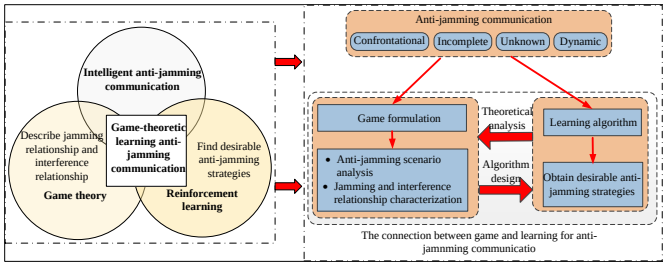

Fig. 6. A framework of game theoretical learning for anti-jamming communication.

As shown in Fig. 6, the game theoretic learning framework is a deep integration of game theory and RL, and it aims to realize the ability of the intelligent anti-jamming communication, in which game theory can accurately describe jamming relationships and mutual interference relationships, and intelligent learning algorithms can obtain desirable anti-jamming strategies through trial and error interactions in dynamic, incomplete and unknown jamming environment. In the game theoretic learning framework, it mainly involves two contents: game formulation and the design of intelligent learning algorithm.

Game formulation:it mainly involves two parts: anti-jamming scenarios analysis and jamming and interference relationships characterization. First, it needs to analyze the anti-jamming scenarios, and different game models have different properties, and can describe different anti-jamming problem. For example, Bayesian game can cope with incomplete information constraints in the anti-jamming field, and Stackelberg game can describe and analyze the sequential behaviors between legitimate users and malicious jammers. Thus, it is important to choose suitable anti-jamming game models according to the properties of the anti-jamming scenarios. Second, it needs to describe two types of relationships, that

TABLE X
SUMMARY OF TRANSFER RL APPROACHES IN ANTI-JAMMING FIELD

Application	Key contribution	Transferred knowledge	Ref.
UAV Relay	Initialize Q-table with the learned experience in similar anti-jamming scenarios for vehicular ad hoc networks to avoid the initialization with an all-zero matrix.	Q-value	[298]
Power control	Initialize Q-table with power control experience in similar scenarios in wireless body area networks.	Q-value	[359]
Channel selection	The learned knowledge from the local regions can be transferred to the neighboring regions.	Q-value	[360]
Channel selection	The learned action and state information can be transferred from a sour task to target task for an actor-critic algorithm in CR network.	Action and state pair	[361]
Power control	Initialize CNN weights with the experience in similar anti-jamming communication scenarios.	CNN weights	[255], [345]
Jiont decision of channel and power	Initialize CNN weights with inter-agent transfer learning to reduce initial random exploration.	CNN weights	[364]
Two-dimensional decision of user mobility and frequency hopping	Initialize CNN weights with the previous experience in similar scenarios in two-dimensional anti-jamming mobile communication system.	CNN weights	[362]
Joint decision of quantization parameters, channel coding rate, modulation type and transmission power	Initialize CNN weights with the learned experience in UAV video transmissions communication.	CNN weights	[363]

is to say, jamming relationships between legitimate users and malicious jammers and mutual interference relationships among legitimate users. For example, mutual interference relationships can be well captured by graphical/hypergraphical game, and jamming relationships can be described by zero-sum game and Bayesian game. A context-aware anti-jamming game framework is shown in Fig.7, and an anti-jamming game model can be chosen according to the properties of anti-jamming scenarios.

Intelligent learning algorithm: It is devoted to obtaining the desirable anti-jamming strategies, which can converge to equilibrium solutions. In anti-jamming problem, it needs to cope with dynamic, incomplete and unknown constraints. Fortunately, intelligent learning technologies are powerful methods to deal with these challenges, and they obtain useful information from feedback due to the interactions with jamming environment. Based on intelligent learning algorithms, the legitimate users can directly or indirectly learn the changing rules of jamming behaviors or jamming environment, and therefore gradually obtain the desirable strategies by adjusting their own behaviors. Some existing intelligent learning algorithms can be found, such as stochastic learning automata [214], and log-linear learning [219], and RL. Due to the powerful learning ability in incomplete and unknown environment, RL will have more applications in practical anti-jamming scenarios in future work.

B. Potential research directions for future investigation

Although a number of studies have been made for anti-jamming communication, this topic still has several unsolved issues for future investigations. Based on the above discussions, this section presents some promising research directions as follows.

1) **Active anti-jamming design**

Although accurate jamming information may be unknown due to the adversarial relations between legitimate users and

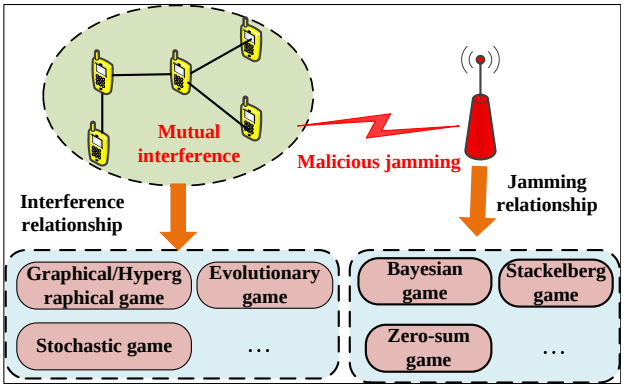

Fig. 7. A framework of context-aware anti-jamming game .

mailicious jammers, the historical jamming information can be observed and utilized. Based on learning mechanisms and historical jamming information, the jamming rules can be discovered and mined. Then, the jamming behaviors can be predicted, and the legitimate users can take actions before the malicious jammers, seize the opportunity and avoid the jamming attacks in advance. In [52], [53], one-step jamming prediction can be realized by Q-learning, and the jamming channel can be avoided with sweeping jamming scenarios. In [101], based on the LSTM model, the multi-step jamming prediction can be designed for sweeping and combing jamming patterns. However, existing active anti-jamming design is preliminary, and they aim at simple jamming scenarios, such as sweeping jamming and comb jamming. More effective active anti-jamming schemes need to be designed in future works.

2) **Collaborative anti-jamming design**

Although it is non-cooperactive between legitimate users and malicious jammers, the collaboration among legitimate users is feasible to improve the anti-jamming capability. Therefore, it is necessary to design multi-level collarorative mechanisms, such as information exchange, multi-domain collaboration and cross-layer collaboration. The information exchange based collaboration mechanism is the most direct mode, and the legitmmate users can share information and experience knowledge to effectively deal with jamming attacks. Most existing anti-jamming schemes were designed relying on a single domain, such as power domain and spectrum domain, and the anti-jamming ability is limited. Therefore, it is an effective mechanism to design anti-jamming schemes from the perspective of multi-domain collaboration, and the anti-jamming performance can be improved by various multi-domain anti-jamming means. Besides, to make full use of the degree of freedom of different layers, such as link layer and network layer, cross-layer anti-jamming design is effective to cope with diversified and sophisticated jamming attacks.

3) **Anti-intelligent jamming design**

With the development of articifical intelligence, intelligent jammers can be produced to pursue more effective jamming patterns, and they can be endowed with some intelligent features, such as learning ability and reconfigurability. They will pose serious threats to wireless communications, and bring great challenges to traditional anti-jamming schemes. Therefore, anti-intelligent jamming design will be important and interesting in future anti-jamming commication. Considering the intelligence of jammers, intelligent countermeasures are inevitable. One basic defense idea is that the legitimate users should have stronger learning ability compared with intelligent jammer in order to obtain better anti-jamming performance. In [341], a DRL-based countermeasure was designed to cope with RL-based intelligent jamming attacks. Moreover, considering the naive weakness of intelligent jammers due to learning ability, the legitimate users can deliberately make some wrong actions to mislead or destroy the learning process of intelligent jammer. Recently, intelligent anti-intelligent jamming mechanisms were explored (e.g. [150], [159]). It would be a promising topic that deserves more attention in anti-jamming field.

4) **Anti-jamming communication for swarm wireless network**

Motivated by biological swarm behaviors, swarm wireless networks have emerged to accomplish complex tasks, such as UAV swarm [321] and aeronautic swarm [140]. A swarm wireless network consists of mulitple swarm nodes, and mutual interference among nodes and external jamming are inevitable. Moreover, there are some obvious characteristics, such as high mobility and dynamic topology. Each node can take on different roles (i.e., transmitter, receiver, and relay). The traditional point-to-point anti-jamming design is difficult to meet the anti-jamming requirements, and cannot support reliable swarm anti-jamming communication. It is, therefore, necessary to make full use of group advantage, and enhance the anti-jamming ability in swarm wirless network.

5) **Anti-jamming communication with hybrid attacks**

In contrast to traditional attackers with a fixed attack mode, intelligent hybrid attackers can cause greater harm to secure communications. Hybrid attackers can select the appropriate multi-attack modes (i.e., [204], [302]), such as eavesdropping and jamming. How to maintain secure and reliable communication transmission under different multi-attack modes is a surely important yet extremely challenging problem. Unlike the single attack mode, multiple attack modes need to be jointly considered in anti-hybrid attacks problem. For instance, if an intelligent hybrid attacker is capable of both eavesdropping and jamming, and a legitimate user only focuses on one attack mode when formulating a defense strategy, the communication security problem remains unsolved. Recently, some preliminary countermeasures were developed in existing studies. In [204], a power domain defense scheme was proposed, and “two birds with one stone strategy” was utilized to simultaneous cope with jamming and eavesdropping. In [302], a RL-based power defense strategy was developed to cope with a smart attack that can choose jamming, eavesdropping and spoofing mode. However, existing anti-hybrid attacks design aim at single attacker scenarios and preliminary power domain defense strategy. In the future, as the number of attackers and types of attacks expand, more effective anti-hybrid attack strategies will be required.

6) **Anti-jamming communication for competing mobile network**

In a competing mobile network [326], [327], the legitimate user and jammer together form a friendly coalition network to

fight against hostile networks. These two opposing networks can be expressed as red team (RT) and blue team (BT). For each network, it includes communicators and jammers, and both attack and defense capabilities are considered. Different from conventional methods, a new confrontation mechanism is formulated, and the jamming and anti-jamming capabilities are jointly designed for each network (i.e., RT or BT). Specifically, the friendly jammers can be employed to deteriorate the hostile communications without weakening own commuications, and legitimate users aim to alleviate mutual interference and combat jamming. In the design of anti-jamming method, it is necessary to jointly consider own friendly jamming, mutual interference and external malicious jamming. Moreover, the hostile users may release some jammed resources due to own friendly jammers, and the legitimate users may obtain some available resources. It is an urgent demand for competing resilient networks in practical tactical wireless networks, and therefore realizes the ability to confront network with network.

VI. Conclusions

In this survey, a comprehensive review is provided for two important anti-jamming solutions, i.e., game theory and reinforcement learning (RL). First, different anti-jamming domains and anti-jamming strategies are disussed, and the technological challenges are globally analyzed from different perspectives. Second, we provide a comprehensive review of each kind of anti-jamming solutions. Specifically, some game theory based anti-jamming solutions are analyzed, such as Bayesian anti-jamming game, Stackelberg anti-jamming game, stochastic anti-jamming game, zero-sum anti-jamming game, graphical/hypergraphical anti-jamming game, and so on. For RL-based anti-jamming solutions, four kinds of RL methods are presented, i.e., Q-learning, multi-armed bandit (MAB), deep reinforcement learning (DRL) and transfer RL. Third, the strengths and limitations are analyzed for each kind of anti-jamming solutions, and some future research directions are discussed. Moreover, each kind of anti-jamming solutions has its strengths and limitations in anti-jamming problem, which implies that the deep integration of game theory and RL will be promising to design effective anti-jamming countermeasures in future works.

References

[1] A. Attar, H. Tang, A. V. Vasilakos, F. R. Yu, and V. C. Leung, "A survey of security challenges in cognitive radio networks: Solutions and future research directions," *Proceedings of the IEEE*, vol. 100, no. 12, pp. 3172–3186, 2012.

[2] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[3] R. K. Sharma and D. B. Rawat, "Advances on security threats and countermeasures for cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1023–1043, 2014.

[4] M. Young and R. Boutaba, "Overcoming adversaries in sensor networks: A survey of theoretical models and algorithmic approaches for tolerating malicious interference," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 4, pp. 617–641, 2011.

[5] A. S. Abdalla, K. Powell, V. Marojevic, and G. Geraci, "Uav-assisted attack prevention, detection, and recovery of 5g networks," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 40–47, 2020.

[6] M. Lichtman, R. P. Jover, M. Labib, R. Rao, V. Marojevic, and J. H. Reed, "Lte/lte-a jamming, spoofing, and sniffing: threat assessment and mitigation," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 54–61, 2016.

[7] M. Lichtman, J. D. Poston, S. Amuru, C. Shahriar, T. C. Clancy, R. M. Buehrer, and J. H. Reed, "A communications jamming taxonomy," *IEEE Security & Privacy*, vol. 14, no. 1, pp. 47–54, 2016.

[8] H. S. D. [S. Amuru and R. M. Buehrer, "On jamming against wireless networks," *IEEE transactions on wireless communications*, vol. 16, no. 1, pp. 412–428, 2017.

[9] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in wsns," *IEEE communications surveys & tutorials*, vol. 11, no. 4, pp. 42–56, 2009.

[10] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, "Jamming games in wireless networks with incomplete information," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 112–118, 2011.

[11] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Communications surveys & tutorials*, vol. 13, no. 2, pp. 245–257, 2010.

[12] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 17, no. 4, pp. 197–215, 2014.

[13] S. Vadlamani, B. Eksioglu, H. Medal, and A. Nandi, "Jamming attacks on wireless networks: A taxonomic survey," *International Journal of Production Economics*, vol. 172, pp. 76–94, 2016.

[14] X. Wang, J. Wang, Y. Xu, J. Chen, L. Jia, X. Liu, and Y. Yang, "Dynamic spectrum anti-jamming communications: Challenges and opportunities," *IEEE communications magazine*, vol. 58, no. 2, pp. 79–85, 2020.

[15] M. A. Aref, S. K. Jayaweera, and E. Yepez, "Survey on cognitive anti-jamming communications," *IET Communications*, vol. 14, no. 18, pp. 3110–3127, 2020.

[16] L. Jia, N. Qi, F. Chu, S. Fang, X. Wang, S. Ma, and S. Feng, "Game-theoretic learning anti-jamming approaches in wireless networks," *IEEE Communications Magazine*, vol. 60, no. 5, pp. 60–66, 2022.

[17] W. Li, J. Chen, X. Liu, X. Wang, Y. Li, D. Liu, and Y. Xu, "Intelligent dynamic spectrum anti-jamming communications: A deep reinforcement learning perspective," *IEEE Wireless Communications*, 2022.

[18] H. Pirayesh and H. Zeng, "Jamming attacks and anti-jamming strategies in wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2022.

[19] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications-a tutorial," *IEEE transactions on Communications*, vol. 30, no. 5, pp. 855–884, 1982.

[20] A. J. Witerbi, "Spread spectrum communications: myths and realities," *The Foundations of the Digital Wireless World: Selected Works of AJ Viterbi*, vol. 2, p. 101, 2010.

[21] C. Popper, M. Strasser, and S. Capkun, "Anti-jamming broadcast communication using uncoordinated spread spectrum techniques," *ieee journal on selected areas in communications*, vol. 28, no. 5, pp. 703–715, 2010.

[22] Q. Wang, P. Xu, K. Ren, and X.-Y. Li, "Towards optimal adaptive ufh-based anti-jamming wireless communication," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 16–30, 2011.

[23] L. Xiao, H. Dai, and P. Ning, "Jamming-resistant collaborative broadcast using uncoordinated frequency hopping," *IEEE transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 297–309, 2011.

[24] A. Cassola, T. Jin, G. Noubir, and B. Thapa, "Efficient spread spectrum communication without preshared secrets," *IEEE Transactions on Mobile Computing*, vol. 12, no. 8, pp. 1669–1680, 2012.

[25] Q. Ling and T. Li, "Message-driven frequency hopping: Design and analysis," *IEEE transactions on wireless communications*, vol. 8, no. 4, pp. 1773–1782, 2009.

[26] L. Zhang, H. Wang, and T. Li, "Anti-jamming message-driven frequency hopping—part i: System design," *IEEE transactions on wireless communications*, vol. 12, no. 1, pp. 70–79, 2012.

[27] L. Zhang and T. Li, "Anti-jamming message-driven frequency hopping—part ii: Capacity analysis under disguised jamming," *IEEE transactions on wireless communications*, vol. 12, no. 1, pp. 80–88, 2012.

[28] H. Wang, L. Zhang, T. Li, and J. Tugnait, "Spectrally efficient jamming mitigation based on code-controlled frequency hopping," *IEEE Transactions on Wireless Communications*, vol. 10, no. 3, pp. 728–732, 2011.

[29] P. Popovski, H. Yomo, and R. Prasad, "Dynamic adaptive frequency hopping for mutually interfering wireless personal area networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 8, pp. 991–1003, 2006.

[30] W. Cheng, Z. Li, F. Gao, L. Liang, and H. Zhang, "Mode hopping for anti-jamming in cognitive radio networks," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 530–535, IEEE, 2018.

[31] L. Liang, W. Cheng, W. Zhang, and H. Zhang, "Mode hopping for anti-jamming in radio vortex wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7018–7032, 2018.

[32] Y. Shi, K. An, and Y. Li, "Index modulation based frequency hopping: Anti-jamming design and analysis," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6930–6942, 2021.

[33] Y. Shi, K. An, X. Lu, and Y. Li, "Enhanced index modulation-based frequency hopping: Resist power-correlated reactive jammer," *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 751–755, 2022.

[34] L. Zhang, Z. Guan, and T. Melodia, "United against the enemy: Anti-jamming based on cross-layer cooperation in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5733–5747, 2016.

[35] L. Zhang, Z. Guan, and T. Melodia, "Cooperative anti-jamming for infrastructure-less wireless networks with stochastic relaying," in *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, pp. 549–557, IEEE, 2014.

[36] F.-T. Hsu and H.-J. Su, "Power allocation strategy against jamming attacks in gaussian fading multichannel," in *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pp. 1062–1066, IEEE, 2014.

[37] S. Ghosh, M. R. Bhatnagar, and B. K. Panigrahi, "Non-cooperative game based defense against broadband jammer in time-critical wireless applications," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 154–159, IEEE, 2016.

[38] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.

[39] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE journal on selected areas in communications*, vol. 23, no. 2, pp. 201–220, 2005.

[40] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game theory in wireless and communication networks*. Cambridge University Press, 2011.

[41] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[42] A. Garnaev, A. Petropulu, W. Trappe, and H. V. Poor, "A multi-jammer game with latency as the user's communication utility," *IEEE Communications Letters*, vol. 24, no. 9, pp. 1899–1903, 2020.

[43] R. H. Gohary, Y. Huang, Z.-Q. Luo, and J.-S. Pang, "A generalized iterative water-filling algorithm for distributed power control in the presence of a jammer," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2660–2674, 2009.

[44] K. Pelechrinis, C. Koufogiannakis, and S. V. Krishnamurthy, "Gaming the jammer: Is frequency hopping effective?," in *2009 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, pp. 1–10, IEEE, 2009.

[45] A. Garnaev and W. Trappe, "Bandwidth scanning when facing interference attacks aimed at reducing spectrum opportunities," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 8, pp. 1916–1930, 2017.

[46] Y. Xu, Y. Xu, X. Dong, G. Ren, J. Chen, X. Wang, L. Jia, and L. Ruan, "Convert harm into benefit: A coordination-learning based dynamic spectrum anti-jamming approach," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13018–13032, 2020.

[47] A. Garnaev, M. Baykal-Gursoy, and H. V. Poor, "A game theoretic analysis of secret and reliable communication with active and passive adversarial modes," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2155–2163, 2015.

[48] Y. Xu, J. Chen, Y. Xu, F. Gu, K. Yao, L. Jia, D. Liu, and X. Wang, "Energy-efficient channel access and data offloading against dynamic jamming attacks," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 4, pp. 1734–1746, 2021.

[49] C. Han, A. Liu, K. An, G. Zheng, and X. Tong, "Distributed uav deployment in hostile environment: A game-theoretic approach," *IEEE Wireless Communications Letters*, vol. 11, no. 1, pp. 126–130, 2021.

[50] B. Bozkurt, A. D. Sezer, S. Gezici, and T. Girici, "A game theoretic approach to channel switching in the presence of jamming," *IEEE Communications Letters*, vol. 25, no. 12, pp. 3927–3931, 2021.

[51] L. Chen and J. Leneutre, "Fight jamming with jamming–a game theoretic analysis of jamming attack in wireless networks and defense strategy," *Computer Networks*, vol. 55, no. 9, pp. 2259–2270, 2011.

[52] F. Slimeni, B. Scheers, Z. Chtourou, and V. Le Nir, "Jamming mitigation in cognitive radio networks using a modified q-learning algorithm," in *2015 International Conference on Military Communications and Information Systems (ICMCIS)*, pp. 1–7, IEEE, 2015.

[53] F. Slimeni, Z. Chtourou, B. Scheers, V. L. Nir, and R. Attia, "Cooperative q-learning based channel selection for cognitive radio networks," *Wireless Networks*, vol. 25, no. 7, pp. 4161–4171, 2019.

[54] L. Xiao, X. Wan, W. Su, Y. Tang, *et al.*, "Anti-jamming underwater transmission with mobility and learning," *IEEE Communications Letters*, vol. 22, no. 3, pp. 542–545, 2018.

[55] Y. Xiao, L. Xiao, Z. Lv, G. Niu, Y. Ding, and W. Xu, "Learning-based low-latency viot video streaming against jamming and interference," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 12–18, 2021.

[56] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pp. 46–57, 2005.

[57] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming sensor networks: attack and defense strategies," *IEEE network*, vol. 20, no. 3, pp. 41–47, 2006.

[58] S. Sciancalepore and R. Di Pietro, "Bittransfer: Mitigating reactive jamming in electronic warfare scenarios," *IEEE Access*, vol. 7, pp. 156175–156190, 2019.

[59] Y. Liang, J. Ren, and T. Li, "Secure ofdm system design and capacity analysis under disguised jamming," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 738–752, 2019.

[60] Q. Peng, P. C. Cosman, and L. B. Milstein, "Spoofing or jamming: Performance analysis of a tactical cognitive radio adversary," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 903–911, 2011.

[61] M. Soysa, P. C. Cosman, and L. B. Milstein, "Optimized spoofing and jamming a cognitive radio," *IEEE Transactions on Communications*, vol. 62, no. 8, pp. 2681–2695, 2014.

[62] S. Amuru and R. M. Buehrer, "Optimal jamming against digital modulation," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2212–2224, 2015.

[63] Q. Liu, M. Li, X. Kong, and N. Zhao, "Disrupting mimo communications with optimal jamming signal design," *IEEE transactions on wireless communications*, vol. 14, no. 10, pp. 5313–5325, 2015.

[64] S. Amuru, C. Tekin, M. van der Schaar, and R. M. Buehrer, "A systematic learning method for optimal jamming," in *2015 IEEE International Conference on Communications (ICC)*, pp. 2822–2827, IEEE, 2015.

[65] S. Amuru, C. Tekin, M. van der Schaar, and R. M. Buehrer, "Jamming bandits—a novel learning method for optimal jamming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2792–2808, 2015.

[66] G. Kim and H. Lim, "Reinforcement learning based beamforming jammer for unknown wireless networks," *IEEE Access*, vol. 8, pp. 210127–210139, 2020.

[67] S. Zhang, H. Tian, X. Chen, Z. Du, L. Huang, Y. Gong, and Y. Xu, "Design and implementation of reinforcement learning-based intelligent jamming system," *IET Communications*, vol. 14, no. 18, pp. 3231–3238, 2020.

[68] F. Wang, M. C. Gursoy, and S. Velipasalar, "Adversarial reinforcement learning in dynamic channel access and power control," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2021.

[69] C. Zhong, F. Wang, M. C. Gursoy, and S. Velipasalar, "Adversarial jamming attacks on deep reinforcement learning based dynamic multichannel access," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2020.

[70] D. Han, A. Li, L. Zhang, Y. Zhang, J. Li, T. Li, T. Zhu, and Y. Zhang, "Deep learning-guided jamming for cross-technology wireless networks: Attack and defense," *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 1922–1932, 2021.

[71] R. Di Pietro and G. Oligeri, "Jamming mitigation in cognitive radio networks," *IEEE Network*, vol. 27, no. 3, pp. 10–15, 2013.

[72] F. M. Aziz, J. S. Shamma, and G. L. Stüber, "Resilience of lte networks against smart jamming attacks," in *2014 IEEE Global Communications Conference*, pp. 734–739, IEEE, 2014.

[73] F. M. Aziz, J. S. Shamma, and G. L. Stüber, "Resilience of lte networks against smart jamming attacks: Wideband model," in *2015 IEEE 26th*

Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 1344–1348, IEEE, 2015.

[74] S. Feng and S. Haykin, "Coordinated cognitive risk control for bridging vehicular radar and communication systems," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[75] S. Feng and S. Haykin, "Cognitive risk control for anti-jamming v2v communications in autonomous vehicle networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9920–9934, 2019.

[76] S. Feng and S. Haykin, "Anti-jamming v2v communication in an integrated uav-cav network with hybrid attackers," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2019.

[77] Z. Lu, W. Wang, and C. Wang, "Modeling, evaluation and detection of jamming attacks in time-critical wireless applications," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1746–1759, 2013.

[78] M. Spuhler, D. Giustiniano, V. Lenders, M. Wilhelm, and J. B. Schmitt, "Detection of reactive jamming in dsss-based wireless communications," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1593–1603, 2014.

[79] R. D. Halloush, "Transmission early-stopping scheme for anti-jamming over delay-sensitive iot applications," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7891–7906, 2019.

[80] S. Xu, W. Xu, C. Pan, and M. Elkashlan, "Detection of jamming attack in non-coherent massive simo systems," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 9, pp. 2387–2399, 2019.

[81] N. An and S. Weber, "Efficiency and detectability of random reactive jamming in carrier sense wireless networks," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 6925–6938, 2019.

[82] M. O. Mughal and S. Kim, "Signal classification and jamming detection in wide-band radios using naïve bayes classifier," *IEEE Communications Letters*, vol. 22, no. 7, pp. 1398–1401, 2018.

[83] A. Benslimane and H. Nguyen-Minh, "Jamming attack model and detection method for beacons under multichannel operation in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6475–6488, 2016.

[84] A. K. Chorppath, T. Alpcan, and H. Boche, "Bayesian mechanisms and detection methods for wireless network with malicious users," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2452–2465, 2015.

[85] N. V. Abhishek and M. Gurusamy, "Jade: Low power jamming detection using machine learning in vehicular networks," *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2210–2214, 2021.

[86] Y. Shi, X. Lu, Y. Niu, and Y. Li, "Efficient jamming identification in wireless communication: Using small sample data driven naive bayes classifier," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1375–1379, 2021.

[87] Y. Cai, K. Shi, F. Song, Y. Xu, X. Wang, and H. Luan, "Jamming pattern recognition using spectrum waterfall: A deep learning method," in *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*, pp. 2113–2117, IEEE, 2019.

[88] F. M. Aziz, J. S. Shamma, and G. L. Stüber, "Jammer-type estimation in lte with a smart jammer repeated game," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7422–7431, 2017.

[89] M. Liu, Z. Liu, W. Lu, Y. Chen, X. Gao, and N. Zhao, "Distributed few-shot learning for intelligent recognition of communication jamming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 395–405, 2021.

[90] X. Wei, Q. Wang, T. Wang, and J. Fan, "Jammer localization in multi-hop wireless network: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 765–799, 2016.

[91] M. A. M. Sadr, M. A. Attari, and R. Amiri, "Robust relay beamforming against jamming attack," *IEEE Communications Letters*, vol. 22, no. 2, pp. 312–315, 2017.

[92] Y. Wu, W. Yang, X. Guan, and Q. Wu, "Uav-enabled relay communication under malicious jamming: Joint trajectory and transmit power optimization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8275–8279, 2021.

[93] Y. Wu, W. Fan, W. Yang, X. Sun, and X. Guan, "Robust trajectory and communication design for multi-uav enabled wireless networks in the presence of jammers," *IEEE Access*, vol. 8, pp. 2893–2905, 2019.

[94] Y. Wu, X. Guan, W. Yang, and Q. Wu, "Uav swarm communication under malicious jamming: joint trajectory and clustering design," *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2264–2268, 2021.

[95] Y. Gao, Y. Wu, Z. Cui, H. Chen, and W. Yang, "Robust design for turning and climbing angle-constrained uav communication under malicious jamming," *IEEE Communications Letters*, vol. 25, no. 2, pp. 584–588, 2020.

[96] Y. Wu, W. Yang, X. Guan, and Q. Wu, "Energy-efficient trajectory design for uav-enabled communication under malicious jamming," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 206–210, 2020.

[97] X. Wang, L. Ming, M. Zhao, and L. Min, "Cooperative anti-jamming strategy and outage probability optimization for multi-hop ad-hoc networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*.

[98] X. Tang, D. Wang, R. Zhang, Z. Chu, and Z. Han, "Jamming mitigation via aerial reconfigurable intelligent surface: Passive beamforming and deployment optimization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6232–6237, 2021.

[99] M. A. M. Sadr, M. Ahmadian-Attari, R. Amiri, and V. V. Sabegh, "Worst-case jamming attack and optimum defense strategy in cooperative relay networks," *IEEE control systems letters*, vol. 3, no. 1, pp. 7–12, 2018.

[100] M.-C. Mah, H.-S. Lim, and A. W.-C. Tan, "Uav relay flight path planning in the presence of jamming signal," *IEEE Access*, vol. 7, pp. 40913–40924, 2019.

[101] Z. Su, N. Qi, L. Jia, J. Chen, Y. Liu, and W. Sun, "End-to-end multi-domain and multi-step jamming prediction in wireless communications," *Electronics Letters*, vol. 57, no. 9, pp. 378–380, 2021.

[102] M. Yang, J. Chen, and Y. Niu, "Dynamic evaluation algorithm for anti-jamming effectiveness of wireless communication equipment based on game theory," in *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pp. 1423–1427, IEEE, 2016.

[103] M. Yang, J. Chen, and Y. Niu, "An evaluation method of anti-jamming capability to communication system based on cloud-evidence theory," in *2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS)*, pp. 1–5, IEEE, 2017.

[104] H. Jung, B. Van Nguyen, I. Song, and K. Kim, "Design of anti-jamming waveforms for time-hopping spread spectrum systems in tone jamming environments," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 728–737, 2019.

[105] D. Yang, J. Zhang, X. Fang, A. Richa, and G. Xue, "Optimal transmission power control in the presence of a smart jammer," in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 5506–5511, IEEE, 2012.

[106] D. Yang, G. Xue, J. Zhang, A. Richa, and X. Fang, "Coping with a smart jammer in wireless networks: A stackelberg game approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4038–4047, 2013.

[107] H. Wang, Y. Fu, R. Song, Z. Shi, and X. Sun, "Power minimization precoding in uplink multi-antenna noma systems with jamming," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 591–602, 2019.

[108] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, "Optimum power allocation for average power constrained jammers in the presence of non-gaussian noise," *IEEE Communications Letters*, vol. 16, no. 8, pp. 1153–1156, 2012.

[109] S. D'Oro, E. Ekici, and S. Palazzo, "Optimal power allocation and scheduling under jamming attacks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 3, pp. 1310–1323, 2016.

[110] K. Pelechrinis, I. Broustis, S. V. Krishnamurthy, and C. Gkantsidis, "A measurement-driven anti-jamming system for 802.11 networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 4, pp. 1208–1222, 2011.

[111] R. El-Bardan, V. Sharma, and P. K. Varshney, "Learning equilibria for power allocation games in cognitive radio networks with a jammer," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1104–1109, IEEE, 2016.

[112] G.-Y. Chang, S.-Y. Wang, and Y.-X. Liu, "A jamming-resistant channel hopping scheme for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6712–6725, 2017.

[113] J.-F. Huang, G.-Y. Chang, and J.-X. Huang, "Anti-jamming rendezvous scheme for cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 648–661, 2016.

[114] H. A. B. Salameh, S. Almajali, M. Ayyash, and H. Elgala, "Spectrum assignment in cognitive radio networks for internet-of-things delay-sensitive applications under jamming attacks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1904–1913, 2018.

[115] S. Luo, S. Zhang, S. Ke, S. Wang, X. Bu, and J. An, "Optimum combining for coherent ffh/ds spread spectrum receivers in the presence of multi-tone jammer," *IEEE Access*, vol. 8, pp. 53097–53106, 2020.

[116] N. Adem, B. Hamdaoui, and A. Yavuz, "Mitigating jamming attacks in mobile cognitive networks through time hopping," *Wireless Communications and Mobile Computing*, vol. 16, no. 17, pp. 3004–3014, 2016.

[117] N. Adem and B. Hamdaoui, "Jamming resiliency and mobility management in cognitive communication networks," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[118] Q. Zhou, Y. Li, and Y. Niu, "A countermeasure against random pulse jamming in time domain based on reinforcement learning," *IEEE access*, vol. 8, pp. 97164–97174, 2020.

[119] Z. Dou, G. Si, Y. Lin, and M. Wang, "An adaptive resource allocation model with anti-jamming in iot network," *IEEE Access*, vol. 7, pp. 93250–93258, 2019.

[120] J. Kim, B. Van Nguyen, H. Jung, and K. Kim, "Th-nrdcsk: A non-coherent time hopping chaotic system for anti-jamming communications.," *IEEE Access*, vol. 7, pp. 144710–144719, 2019.

[121] M. Tiloca, D. De Guglielmo, G. Dini, G. Anastasi, and S. K. Das, "Jammy: a distributed and dynamic solution to selective jamming attack in tdma wsns," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 392–405, 2015.

[122] Q. Yan, H. Zeng, T. Jiang, M. Li, W. Lou, and Y. T. Hou, "Jamming resilient communication using mimo interference cancellation," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 7, pp. 1486–1499, 2016.

[123] T. T. Do, E. Björnson, E. G. Larsson, and S. M. Razavizadeh, "r123,," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 210–223, 2017.

[124] H. Akhlaghpasand, E. Björnson, and S. M. Razavizadeh, "Jamming-robust uplink transmission for spatially correlated massive mimo systems," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3495–3504, 2020.

[125] H. Pirayesh, P. K. Sangdeh, and H. Zeng, "Securing zigbee communications against constant jamming attack using neural network," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4957–4968, 2020.

[126] X. He, H. Dai, and P. Ning, "Dynamic adaptive anti-jamming via controlled mobility," *IEEE transactions on wireless communications*, vol. 13, no. 8, pp. 4374–4388, 2014.

[127] X. He, H. Dai, and P. Ning, "Dynamic adaptive anti-jamming via controlled mobility," *IEEE transactions on wireless communications*, vol. 13, no. 8, pp. 4374–4388, 2014.

[128] X. He, H. Dai, and P. Ning, "[ieee 2013 ieee conference on communications and network security (cns) - national harbor, md, usa (2013.10.14-2013.10.16)] 2013 ieee conference on communications and network security (cns) - dynamic adaptive anti-jamming via controlled mobility," pp. 1–9, 2013.

[129] P. Gu, C. Hua, R. Khatoun, Y. Wu, and A. Serhrouchni, "Cooperative antijamming relaying for control channel jamming in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7033–7046, 2018.

[130] P. Gu, C. Hua, W. Xu, R. Khatoun, Y. Wu, and A. Serhrouchni, "Control channel anti-jamming in vehicular networks via cooperative relay beamforming," *IEEE internet of things journal*, vol. 7, no. 6, pp. 5064–5077, 2020.

[131] P. Gu, C. Hua, R. Khatoun, Y. Wu, and A. Serhrouchni, "Cooperative antijamming relaying for control channel jamming in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7033–7046, 2018.

[132] X. Li, Y. Zhu, and B. Li, "Optimal anti-jamming strategy in sensor networks," in *2012 IEEE International Conference on Communications (ICC)*, pp. 178–182, IEEE, 2012.

[133] K. Wu, P. C. Cosman, and L. B. Milstein, "Multicarrier ds-cdma system under fast rician fading and partial-time partial-band jamming," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7183–7194, 2019.

[134] M. R. Masse, M. B. Pursley, and J. S. Skinner, "Adaptive coding for frequency-hop transmission over fading channels with partial-band interference," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 854–862, 2011.

[135] M. B. Pursley and J. S. Skinner, "Adaptive coding for frequency-hop transmission in mobile ad hoc networks with partial-band interference," *IEEE Transactions on Communications*, vol. 57, no. 3, pp. 801–811, 2009.

[136] K. Xu, Q. Wang, and K. Ren, *Joint UFH and power control for effective wireless anti-jamming communication*. IEEE, 2012.

[137] K. Dabcevic, A. Betancourt, L. Marcenaro, and C. S. Regazzoni, "A fictitious play-based game-theoretical approach to alleviating jamming attacks for cognitive radios," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8158–8162, IEEE, 2014.

[138] Q. Wang and M. Liu, "Joint control of transmission power and channel switching against adaptive jamming," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 909–916, IEEE, 2013.

[139] D. B. Rawat and M. Song, "Securing space communication systems against reactive cognitive jammer," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1428–1433, IEEE, 2015.

[140] H. Li, J. Luo, and C. Liu, "Selfish bandit-based cognitive anti-jamming strategy for aeronautic swarm network in presence of multiple jammer," *IEEE Access*, vol. 7, pp. 30234–30243, 2019.

[141] A. Alipour-Fanid, M. Dabaghchian, N. Wang, L. Jiao, and K. Zeng, "Online-learning-based defense against jamming attacks in multichannel wireless cps," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13278–13290, 2021.

[142] L. Jia, Y. Xu, Y. Sun, S. Feng, L. Yu, and A. Anpalagan, "A multi-domain anti-jamming defense scheme in heterogeneous wireless networks," *IEEE Access*, vol. 6, pp. 40177–40188, 2018.

[143] Y. Li, S. Bai, and Z. Gao, "A multi-domain anti-jamming strategy using stackelberg game in wireless relay networks," *IEEE Access*, vol. 8, pp. 173609–173617, 2020.

[144] X. Pei, X. Wang, J. Yao, C. Yao, J. Ge, L. Huang, and D. Liu, "Joint time-frequency anti-jamming communications: a reinforcement learning approach," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, IEEE, 2019.

[145] P. Bhavathankar, S. Chatterjee, and S. Misra, "Link-quality aware path selection in the presence of proactive jamming in fallible wireless sensor networks," *IEEE transactions on communications*, vol. 66, no. 4, pp. 1689–1704, 2017.

[146] I. K. Azogu, M. T. Ferreira, J. A. Larcom, and H. Liu, "A new anti-jamming strategy for vanet metrics-directed security defense," in *2013 IEEE Globecom Workshops (GC Wkshps)*, pp. 1344–1349, IEEE, 2013.

[147] J. Song, Q. Zhang, S. Kadhe, M. Bakshi, and S. Jaggi, "Stealthy communication over adversarially jammed multipath networks," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7473–7484, 2020.

[148] S. Nan, S. Brahma, C. A. Kamhoua, and N. O. Leslie, "Mitigation of jamming attacks via deception," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, IEEE, 2020.

[149] Y. Shi, Y. E. Sagduyu, T. Erpek, K. Davaslioglu, Z. Lu, and J. H. Li, "Adversarial deep learning for cognitive radio security: Jamming attack and defense strategies," in *2018 IEEE international conference on communications workshops (ICC Workshops)*, pp. 1–6, IEEE, 2018.

[150] T. Erpek, Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 1, pp. 2–14, 2018.

[151] D. T. Hoang, D. N. Nguyen, M. A. Alsheikh, S. Gong, E. Dutkiewicz, D. Niyato, and Z. Han, "" borrowing arrows with thatched boats": The art of defeating reactive jammers in iot networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 79–87, 2020.

[152] N. Van Huynh, D. T. Hoang, D. N. Nguyen, and E. Dutkiewicz, "Deepfake: Deep dueling-based deception strategy to defeat reactive jammers," *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6898–6914, 2021.

[153] N. Van Huynh, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, and M. Mueck, "Defeating smart and reactive jammers with unlimited power," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2020.

[154] S.-H. Lim, S. Han, J. Lee, Y. Eun, and J.-W. Choi, "Decoy signal based strategic beamforming against high-power reactive jamming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10054–10058, 2018.

[155] A. Pourranjbar, G. Kaddoum, A. Ferdowsi, and W. Saad, "Reinforcement learning for deceiving reactive jammers in wireless networks," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3682–3697, 2021.

[156] S. Bhunia, E. Miles, S. Sengupta, and F. Vázquez-Abad, "Cr-honeynet: A cognitive radio learning and decoy-based sustenance mechanism to avoid intelligent jammer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 567–581, 2018.

[157] S. Bhunia, S. Sengupta, and F. Vázquez-Abad, "Cr-honeynet: A learning & decoy based sustenance mechanism against jamming attack in crn," in *2014 IEEE Military Communications Conference*, pp. 1173–1180, IEEE, 2014.

[158] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, "Performance analysis of wireless energy harvesting cognitive radio networks under smart jamming attacks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 2, pp. 200–216, 2015.

[159] W. Li, J. Wang, L. Li, X. Chen, W. Huang, and S. Li, "Countermeasure for smart jamming threat: A deceptively adversarial attack approach," in *ICC 2021-IEEE International Conference on Communications*, pp. 1–6, IEEE, 2021.

[160] A. Sheikholeslami, M. Ghaderi, H. Pishro-Nik, and D. Goeckel, "Energy-efficient routing in wireless networks in the presence of jamming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6828–6842, 2016.

[161] A. Sheikholeslami, M. Ghaderi, H. Pishro-Nik, and D. Goeckel, "Jamming-aware minimum energy routing in wireless networks," in *2014 IEEE International Conference on Communications (ICC)*, pp. 2313–2318, IEEE, 2014.

[162] P. Tague, S. Nabar, J. A. Ritcey, and R. Poovendran, "Jamming-aware traffic allocation for multiple-path routing using portfolio selection," *IEEE/ACM Transactions On Networking*, vol. 19, no. 1, pp. 184–194, 2010.

[163] H. Mustafa, X. Zhang, Z. Liu, W. Xu, and A. Perrig, "Jamming-resilient multipath routing," *IEEE transactions on dependable and secure computing*, vol. 9, no. 6, pp. 852–864, 2012.

[164] M. Balakrishnan, H. Huang, R. Asorey-Cacheda, S. Misra, S. Pawar, and Y. Jaradat, "Measures and countermeasures for null frequency jamming of on-demand routing protocols in wireless ad hoc networks," *IEEE transactions on wireless communications*, vol. 11, no. 11, pp. 3860–3868, 2012.

[165] Y. Xuan, Y. Shen, N. P. Nguyen, and M. T. Thai, "A trigger identification service for defending reactive jammers in wsn," *IEEE Transactions on Mobile Computing*, vol. 11, no. 5, pp. 793–806, 2011.

[166] Y. Z. Jembre and Y.-J. Choi, "Distributed and jamming-resistant channel assignment and routing for multi-hop wireless networks," *IEEE Access*, vol. 6, pp. 76402–76415, 2018.

[167] Z. H. Abbas, G. Abbas, M. S. Haroon, and F. Muhammad, "Analysis of interference management in heterogeneous cellular networks in the presence of wideband jammers," *IEEE Communications Letters*, vol. 24, no. 5, pp. 1138–1141, 2020.

[168] F. Muhammad, M. S. Haroon, Z. H. Abbas, G. Abbas, and S. Kim, "Uplink interference management for hetnets stressed by clustered wide-band jammers," *IEEE access*, vol. 7, pp. 182679–182690, 2019.

[169] S.-Y. Chang, Y.-C. Hu, and N. Laurenti, "Simplemac: A simple wireless mac-layer countermeasure to intelligent and insider jammers," *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 1095–1108, 2015.

[170] N. Zhao, F. R. Yu, M. Li, and V. C. Leung, "Anti-eavesdropping schemes for interference alignment (ia)-based wireless networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5719–5732, 2016.

[171] J. Guo, N. Zhao, F. R. Yu, X. Liu, and V. C. Leung, "Exploiting adversarial jamming signals for energy harvesting in interference networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1267–1280, 2016.

[172] J. Guo, N. Zhao, Z. Yang, F. R. Yu, Y. Chen, and V. C. Leung, "Proactive jamming toward interference alignment networks: Beneficial and adversarial aspects," *IEEE Systems Journal*, vol. 13, no. 1, pp. 412–423, 2017.

[173] W. Li, J. Wang, L. Li, G. Zhang, Z. Dang, and S. Li, "Intelligent anti-jamming communication with continuous action decision for ultra-dense network," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–7, IEEE, 2019.

[174] J. Farah, E. P. Simon, P. Laly, and G. Delbarre, "Efficient combinations of noma with distributed antenna systems based on channel measurements for mitigating jamming attacks," *IEEE Systems Journal*, vol. 15, no. 2, pp. 2212–2221, 2020.

[175] A. Garnaev, A. P. Petropulu, W. Trappe, and H. V. Poor, "A jamming game with rival-type uncertainty," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5359–5372, 2020.

[176] A. Garnaev, Y. Liu, and W. Trappe, "Anti-jamming strategy versus a low-power jamming attack when intelligence of adversary's attack type is unknown," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp. 49–56, 2015.

[177] R. El-Bardan, S. Brahma, and P. K. Varshney, "Strategic power allocation with incomplete information in the presence of a jammer," *IEEE Transactions on Communications*, vol. 64, no. 8, pp. 3467–3479, 2016.

[178] M. Scalabrin, V. Vadori, A. V. Guglielmi, and L. Badia, "A zero-sum jamming game with incomplete position information in wireless scenarios," in *Proceedings of European Wireless 2015; 21th European Wireless Conference*, pp. 1–6, VDE, 2015.

[179] V. Vadori, M. Scalabrin, A. V. Guglielmi, and L. Badia, "Jamming in underwater sensor networks as a bayesian zero-sum game with position uncertainty," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2015.

[180] X. Liu, Y. Xu, L. Jia, Q. Wu, and A. Anpalagan, "Anti-jamming communications using spectrum waterfall: A deep reinforcement learning approach," *IEEE Communications Letters*, vol. 22, no. 5, pp. 998–1001, 2018.

[181] Y. Li, Y. Xu, Y. Xu, X. Liu, X. Wang, W. Li, and A. Anpalagan, "Dynamic spectrum anti-jamming in broadband communications: A hierarchical deep reinforcement learning approach," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1616–1619, 2020.

[182] P. Zhou and T. Jiang, "Toward optimal adaptive wireless communications in unknown environments," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3655–3667, 2016.

[183] F. Yao, L. Jia, Y. Sun, Y. Xu, S. Feng, and Y. Zhu, "A hierarchical learning approach to anti-jamming channel selection strategies," *Wireless Networks*, vol. 25, no. 1, pp. 201–213, 2019.

[184] L. Jia, Y. Xu, Y. Sun, S. Feng, L. Yu, and A. Anpalagan, "A game-theoretic learning approach for anti-jamming dynamic spectrum access in dense wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1646–1656, 2018.

[185] F. Yao and L. Jia, "A collaborative multi-agent reinforcement learning anti-jamming algorithm in wireless networks," *IEEE wireless communications letters*, vol. 8, no. 4, pp. 1024–1027, 2019.

[186] Y. Zhang, L. Jia, N. Qi, Y. Xu, and X. Chen, "A multi-agent reinforcement learning anti-jamming method with partially overlapping channels," *IET Communications*, vol. 15, no. 19, pp. 2461–2468, 2021.

[187] Q. Zhou, Y. Li, and Y. Niu, "Intelligent anti-jamming communication for wireless sensor networks: A multi-agent reinforcement learning approach," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 775–784, 2021.

[188] I. Elleuch, A. Pourranjbar, and G. Kaddoum, "A novel distributed multi-agent reinforcement learning algorithm against jamming attacks," *IEEE Communications Letters*, vol. 25, no. 10, pp. 3204–3208, 2021.

[189] Y. E. Sagduyu, R. Berry, and A. Ephremides, "Mac games for distributed wireless network security with incomplete information of selfish and malicious user types," in *2009 International Conference on Game Theory for Networks*, pp. 130–139, IEEE, 2009.

[190] A. Garnaev, W. Trappe, and A. Petropulu, "Combating jamming in wireless networks: A bayesian game with jammer's channel uncertainty," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2447–2451, IEEE, 2019.

[191] Y. Xu, G. Ren, J. Chen, Y. Luo, L. Jia, X. Liu, Y. Yang, and Y. Xu, "A one-leader multi-follower bayesian-stackelberg game for anti-jamming transmission in uav communication networks," *Ieee Access*, vol. 6, pp. 21697–21709, 2018.

[192] M. J. Abdel-Rahman and M. Krunz, "Game-theoretic quorum-based frequency hopping for anti-jamming rendezvous in dsa networks," in *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, pp. 248–258, IEEE, 2014.

[193] R. El-Bardan, S. Brahma, and P. K. Varshney, "Power control with jammer location uncertainty: A game theoretic perspective," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, IEEE, 2014.

[194] A. Garnaev, A. Petropulu, W. Trappe, and H. V. Poor, "A power control game with uncertainty on the type of the jammer," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–5, IEEE, 2019.

[195] A. Garnaev and W. Trappe, "Fair resource allocation under an unknown jamming attack: a bayesian game," in *2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 227–232, IEEE, 2014.

[196] E. Altman, K. Avrachenkov, and A. Garnaev, "Jamming game with incomplete information about the jammer," in *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools*, pp. 1–9, 2009.

[197] E. Altman, K. Avrachenkov, and A. Garnaev, "Jamming in wireless networks under uncertainty," *Mobile Networks and Applications*, vol. 16, no. 2, pp. 246–254, 2011.

[198] A. Garnaev, Y. Hayel, and E. Altman, "A bayesian jamming game in an ofdm wireless network," in *2012 10th International Symposium on*

Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), pp. 41–48, IEEE, 2012.

[199] L. Jia, F. Yao, Y. Sun, Y. Niu, and Y. Zhu, "Bayesian stackelberg game for antijamming transmission with incomplete information," *IEEE Communications Letters*, vol. 20, no. 10, pp. 1991–1994, 2016.

[200] Z. Feng, G. Ren, J. Chen, X. Zhang, Y. Luo, M. Wang, and Y. Xu, "Power control in relay-assisted anti-jamming systems: A bayesian three-layer stackelberg game approach," *IEEE Access*, vol. 7, pp. 14623–14636, 2019.

[201] X. Tang, P. Ren, Y. Wang, Q. Du, and L. Sun, "Securing wireless transmission against reactive jamming: A stackelberg game framework," in *2015 IEEE global communications conference (GLOBECOM)*, pp. 1–6, IEEE, 2015.

[202] Y. Xu, G. Ren, J. Chen, X. Zhang, L. Jia, Z. Feng, and Y. Xu, "Joint power and trajectory optimization in uav anti-jamming communication networks," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–5, IEEE, 2019.

[203] Z. Su, N. Qi, Y. Yan, Z. Du, J. Chen, Z. Feng, and Q. Wu, "Guarding legal communication with smart jammer: Stackelberg game based power control analysis," *China Communications*, vol. 18, no. 4, pp. 126–136, 2021.

[204] N. Qi, W. Wang, F. Zhou, L. Jia, Q. Wu, S. Jin, and M. Xiao, "Two birds with one stone: Simultaneous jamming and eavesdropping with the bayesian-stackelberg game," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8013–8027, 2021.

[205] L. Xiao, T. Chen, J. Liu, and H. Dai, "Anti-jamming transmission stackelberg game with observation errors," *IEEE communications letters*, vol. 19, no. 6, pp. 949–952, 2015.

[206] L. Zhao, H. Xu, J. Zhang, and H. Yang, "Resilient control for wireless cyber-physical systems subject to jamming attacks: A cross-layer dynamic game approach," *IEEE Transactions on Cybernetics*, 2020.

[207] H. Yang, M. Shi, Y. Xia, and P. Zhang, "Security research on wireless networked control systems subject to jamming attacks," *IEEE transactions on cybernetics*, vol. 49, no. 6, pp. 2022–2031, 2018.

[208] X. Tang, P. Ren, and Z. Han, "Jamming mitigation via hierarchical security game for iot communications," *IEEE Access*, vol. 6, pp. 5766–5779, 2018.

[209] S. G. Hong, Y. M. Hwang, S. Y. Lee, Y. Shin, D. I. Kim, and J. Y. Kim, "Game-theoretic modeling of backscatter wireless sensor networks under smart interference," *IEEE Communications Letters*, vol. 22, no. 4, pp. 804–807, 2017.

[210] A. Garnaev, A. Petropulu, W. Trappe, and H. V. Poor, "A multi-jammer power control game," *IEEE Communications Letters*, vol. 25, no. 9, pp. 3031–3035, 2021.

[211] Y. Li, L. Xiao, J. Liu, and Y. Tang, "Power control stackelberg game in cooperative anti-jamming communications," in *The 2014 5th International Conference on Game Theory for Networks*, pp. 1–6, IEEE, 2014.

[212] L. Xiao, Y. Li, J. Liu, and Y. Zhao, "Power control with reinforcement learning in cooperative cognitive radio networks against jamming," *The Journal of Supercomputing*, vol. 71, no. 9, pp. 3237–3257, 2015.

[213] S. Lv, L. Xiao, Q. Hu, X. Wang, C. Hu, and L. Sun, "Anti-jamming power control game in unmanned aerial vehicle networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2017.

[214] L. Jia, F. Yao, Y. Sun, Y. Xu, S. Feng, and A. Anpalagan, "A hierarchical learning solution for anti-jamming stackelberg game with discrete power strategies," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 818–821, 2017.

[215] N. Gao, Z. Qin, X. Jing, Q. Ni, and S. Jin, "Anti-intelligent uav jamming strategy via deep q-networks," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 569–581, 2019.

[216] F. Slimeni, V. Le Nir, B. Scheers, Z. Chtourou, and R. Attia, "Optimal power allocation over parallel gaussian channels in cognitive radio and jammer games," *Iet Communications*, vol. 10, no. 8, pp. 980–986, 2016.

[217] L. Yu, Q. Wu, Y. Xu, G. Ding, and L. Jia, "Power control games for multi-user anti-jamming communications," *Wireless Networks*, vol. 25, no. 5, pp. 2365–2374, 2019.

[218] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, "A mobile offloading game against smart attacks," *IEEE Access*, vol. 4, pp. 2281–2291, 2016.

[219] N. Qi, W. Wang, M. Xiao, L. Jia, S. Jin, Q. Zhu, and T. A. Tsiftsis, "A learning-based spectrum access stackelberg game: Friendly jammer-assisted communication confrontation," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 700–713, 2021.

[220] L. Xiao, Y. Li, C. Dai, H. Dai, and H. V. Poor, "Reinforcement learning-based noma power allocation in the presence of smart jamming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3377–3389, 2017.

[221] J. Liu, X. Wang, S. Shen, Z. Fang, S. Yu, G. Yue, and M. Li, "Intelligent jamming defense using dnn stackelberg game in sensor edge cloud," *IEEE Internet of Things Journal*, vol. 9, no. 6, pp. 4356–4370, 2021.

[222] X. Zhang, H. Wang, Y. Xu, Z. Feng, and Y. Zhang, "Put others before itself: A multi-leader one-follower anti-jamming stackelberg game against tracking jammer," *China Communications*, vol. 18, no. 11, pp. 168–181, 2021.

[223] Y. Zhang, Y. Xu, Y. Xu, Y. Yang, Y. Luo, Q. Wu, and X. Liu, "A multi-leader one-follower stackelberg game approach for cooperative anti-jamming: no pains, no gains," *IEEE Communications Letters*, vol. 22, no. 8, pp. 1680–1683, 2018.

[224] C. Han and Y. Niu, "Cross-layer anti-jamming scheme: A hierarchical learning approach," *IEEE Access*, vol. 6, pp. 34874–34883, 2018.

[225] C. Han, A. Liu, H. Wang, L. Huo, and X. Liang, "Dynamic anti-jamming coalition for satellite-enabled army iot: A distributed game approach," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 10932–10944, 2020.

[226] C. Han, A. Liu, L. Huo, H. Wang, and X. Liang, "Anti-jamming routing for internet of satellites: a reinforcement learning approach," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2877–2881, IEEE, 2020.

[227] C. Han, L. Huo, X. Tong, H. Wang, and X. Liu, "Spatial anti-jamming scheme for internet of satellites based on the deep reinforcement learning and stackelberg game," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5331–5342, 2020.

[228] I. K. Ahmed and A. O. Fapojuwo, "Stackelberg equilibria of an anti-jamming game in cooperative cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 1, pp. 121–134, 2017.

[229] X. Wang, X. Chen, M. Wang, and S. Dong, "Decentralized reinforcement learning based anti-jamming communication for self-organizing networks," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2021.

[230] X. Wang, Y. Xu, J. Chen, C. Li, X. Liu, D. Liu, and Y. Xu, "Mean field reinforcement learning based anti-jamming communications for ultra-dense internet of things in 6g," in *2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 195–200, IEEE, 2020.

[231] B. Wang, Y. Wu, K. R. Liu, and T. C. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE journal on selected areas in communications*, vol. 29, no. 4, pp. 877–889, 2011.

[232] X. He, H. Dai, and P. Ning, "Improving learning and adaptation in security games by exploiting information asymmetry," in *2015 IEEE conference on computer communications (INFOCOM)*, pp. 1787–1795, IEEE, 2015.

[233] X. He, H. Dai, and P. Ning, "Faster learning and adaptation in security games by exploiting information asymmetry," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3429–3443, 2016.

[234] K. Ibrahim, S. X. Ng, I. M. Qureshi, A. N. Malik, and S. Muhaidat, "Anti-jamming game to combat intelligent jamming for cognitive radio networks," *IEEE Access*, vol. 9, pp. 137941–137956, 2021.

[235] K. Ibrahim, I. M. Qureshi, A. N. Malik, and S. X. Ng, "Bandwidth-efficient frequency hopping based anti-jamming game for cognitive radio assisted wireless sensor networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pp. 1–5, IEEE, 2021.

[236] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 247–254, IEEE, 2014.

[237] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless systems," *IEEE Transactions on Mobile Computing*, vol. 15, no. 9, pp. 2247–2259, 2015.

[238] B. DeBruhl, C. Kroer, A. Datta, T. Sandholm, and P. Tague, "Power napping with loud neighbors: optimal energy-constrained jamming and anti-jamming," in *Proceedings of the 2014 ACM conference on Security and privacy in wireless & mobile networks*, pp. 117–128, 2014.

[239] T. Song, W. E. Stark, T. Li, and J. K. Tugnait, "Optimal multiband transmission under hostile jamming," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 4013–4027, 2016.

[240] Q. Wang, T. Nguyen, K. Pham, and H. Kwon, "Mitigating jamming attack: A game-theoretic perspective," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6063–6074, 2018.

[241] F. Slimeni, B. Scheers, V. Le Nir, Z. Chtourou, and R. Attia, "Learning multi-channel power allocation against smart jammer in cognitive radio networks," in *2016 International Conference on Military Communications and Information Systems (ICMCIS)*, pp. 1–7, IEEE, 2016.

[242] K. Pelechrinis, C. Koufogiannakis, and S. V. Krishnamurthy, "On the efficacy of frequency hopping in coping with jamming attacks in 802.11 networks," *IEEE transactions on wireless communications*, vol. 9, no. 10, pp. 3258–3271, 2010.

[243] Y. Xu, Y. Xu, G. Ren, J. Chen, C. Yao, L. Jia, and D. Liu, "Context-aware coordinated anti-jamming communications: A multi-pattern stochastic learning approach," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–5, IEEE, 2021.

[244] Y. Xu, Y. Xu, G. Ren, J. Chen, C. Yao, L. Jia, D. Liu, and X. Wang, "Play it by ear: Context-aware distributed coordinated anti-jamming channel access," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 5279–5293, 2021.

[245] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan, "Stackelberg game approaches for anti-jamming defence in wireless networks," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 120–128, 2018.

[246] A. Garnaev, M. Baykal-Gursoy, and H. V. Poor, "Security games with unknown adversarial strategies," *IEEE transactions on cybernetics*, vol. 46, no. 10, pp. 2291–2299, 2015.

[247] A. Garnaev and W. Trappe, "A bandwidth monitoring strategy under uncertainty of the adversary's activity," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 837–849, 2015.

[248] A. Garnaev and W. Trappe, "The rival might be not smart: revising a cdma jamming game," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2018.

[249] Y. E. Sagduyu, R. A. Berryt, and A. Ephremidesi, "Wireless jamming attacks under dynamic traffic uncertainty," in *8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, pp. 303–312, IEEE, 2010.

[250] G. Li, Z. He, C. Xing, C. Chen, and Q. Liao, "Qos-based anti-jamming algorithm design for distributed wireless networks," in *2013 International Conference on Wireless Communications and Signal Processing*, pp. 1–5, IEEE, 2013.

[251] A. Garnaev, W. Trappe, and A. Petropulu, "Equilibrium strategies for an ofdm network that might be under a jamming attack," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, IEEE, 2017.

[252] Z. Shen, K. Xu, and X. Xia, "Beam-domain anti-jamming transmission for downlink massive mimo systems: A stackelberg game perspective," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2727–2742, 2021.

[253] A. Garnaev, A. Petropulu, W. Trappe, and H. V. Poor, "An anti-jamming multiple access channel game using latency as metric," *IEEE Wireless Communications Letters*, 2022.

[254] K. Liu, P. Li, C. Liu, L. Xiao, and L. Jia, "Uav-aided anti-jamming maritime communications: a deep reinforcement learning approach," in *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, IEEE, 2021.

[255] J. Lu, D. He, and Z. Wang, "Deepantijam: Stackelberg game-oriented secure transmission via deep reinforcement learning," *IEEE Communications Letters*, 2022.

[256] S. Misra, A. Mondal, P. Bhavathankar, and M.-S. Alouini, "M-jaw: Mobility-based jamming avoidance in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5381–5390, 2020.

[257] S. d'Oro, L. Galluccio, G. Morabito, S. Palazzo, L. Chen, and F. Martignon, "Defeating jamming with the power of silence: A game-theoretic analysis," *IEEE transactions on wireless communications*, vol. 14, no. 5, pp. 2337–2352, 2014.

[258] A. Pourranjbar, G. Kaddoum, and K. Aghababaiyan, "Deceiving-based anti-jamming against single-tone and multitone reactive jammers," *IEEE Transactions on Communications*, vol. 70, no. 9, pp. 6133–6148, 2022.

[259] Y. Sun, Y. Zhu, K. An, G. Zheng, S. Chatzinotas, K.-k. Wong, and P. Liu, "Robust design for ris-assisted anti-jamming communications with imperfect angular information: A game-theoretic perspective," *IEEE Transactions on Vehicular Technology*, 2022.

[260] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156–172, 2008.

[261] M. A. Aref, S. K. Jayaweera, and S. Machuzak, "Multi-agent reinforcement learning based cognitive anti-jamming," in *2017 IEEE wireless communications and networking conference (WCNC)*, pp. 1–6, IEEE, 2017.

[262] M. A. Aref and S. K. Jayaweera, "A novel cognitive anti-jamming stochastic game," in *2017 Cognitive Communications for Aerospace Applications Workshop (CCAA)*, pp. 1–4, IEEE, 2017.

[263] M. A. Aref and S. K. Jayaweera, "A cognitive anti-jamming and interference-avoidance stochastic game," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, pp. 520–527, IEEE, 2017.

[264] M. A. Aref and S. K. Jayaweera, "Jamming-resilient wideband cognitive radios with multi-agent reinforcement learning," *International Journal of Software Science and Computational Intelligence (IJSSCI)*, vol. 10, no. 3, pp. 1–23, 2018.

[265] M. G. Oskoui, P. Khorramshahi, and J. A. Salehi, "Using game theory to battle jammer in control channels of cognitive radio ad hoc networks," in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–5, IEEE, 2016.

[266] Z. Yin, Y. Lin, Y. Zhang, Y. Qian, F. Shu, and J. Li, "Collaborative multi-agent reinforcement learning aided resource allocation for uav anti-jamming communication," *IEEE Internet of Things Journal*, 2022.

[267] Q. Zhu, H. Li, Z. Han, and T. Başar, "A stochastic game model for jamming in multi-channel cognitive radio systems," in *2010 IEEE International Conference on Communications*, pp. 1–6, IEEE, 2010.

[268] H. Noori and S. Sadeghi Vilni, "Jamming and anti-jamming in interference channels: a stochastic game approach," *IET Communications*, vol. 14, no. 4, pp. 682–692, 2020.

[269] M. A. Alavijeh, B. Maham, Z. Han, and S. Nader-Esfahani, "Efficient anti-jamming truthful spectrum auction among secondary users in cognitive radio networks," in *2013 IEEE International Conference on Communications (ICC)*, pp. 2812–2816, IEEE, 2013.

[270] J. He, C. Chen, S. Zhu, B. Yang, and X. Guan, "Antijamming game framework for secure state estimation in power systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2628–2637, 2018.

[271] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Cognitive networks with in-band full-duplex radios: Jamming attacks and countermeasures," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 296–309, 2019.

[272] B. F. Lo and I. F. Akyildiz, "Multiagent jamming-resilient control channel game for cognitive radio ad hoc networks," in *2012 IEEE International Conference on Communications (ICC)*, pp. 1821–1826, IEEE, 2012.

[273] D. Niyato, P. Wang, D. I. Kim, Z. Han, and L. Xiao, "Game theoretic modeling of jamming attack in wireless powered communication networks," in *2015 IEEE International Conference on Communications (ICC)*, pp. 6018–6023, IEEE, 2015.

[274] X. Zhou, D. Niyato, and A. Hjorungnes, "Optimizing training-based transmission against smart jamming," *IEEE transactions on vehicular technology*, vol. 60, no. 6, pp. 2644–2655, 2011.

[275] C. Chen, M. Song, C. Xin, and J. Backens, "A game-theoretical anti-jamming scheme for cognitive radio networks," *IEEE network*, vol. 27, no. 3, pp. 22–27, 2013.

[276] S. Wei, R. Kannan, V. Chakravarthy, and M. Rangaswamy, "Csi usage over parallel fading channels under jamming attacks: A game theory study," *IEEE transactions on communications*, vol. 60, no. 4, pp. 1167–1175, 2012.

[277] G. Rezgui, E. V. Belmega, and A. Chorti, "Mitigating jamming attacks using energy harvesting," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 297–300, 2018.

[278] A. Signori, F. Chiariotti, F. Campagnaro, and M. Zorzi, "A game-theoretic and experimental analysis of energy-depleting underwater jamming attacks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9793–9804, 2020.

[279] S. T. Jose and A. A. Kulkarni, "Shannon meets von neumann: A minimax theorem for channel coding in the presence of a jammer," *IEEE Transactions on Information Theory*, vol. 66, no. 5, pp. 2842–2859, 2020.

[280] K. Firouzbakht, G. Noubir, and M. Salehi, "On the performance of adaptive packetized wireless communication links under jamming," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3481–3495, 2014.

[281] A. Signori, F. Chiariotti, F. Campagnaro, R. Petroccia, K. Pelekanakis, P. Paglierani, J. Alves, and M. Zorzi, "A geometry-based game theoretical model of blind and reactive underwater jamming," *IEEE Transactions on Wireless Communications*, 2021.

[282] Y. Xu, Q. Wu, L. Shen, J. Wang, and A. Anpalagan, "Opportunistic spectrum access with spatial reuse: Graphical game and uncoupled

learning solutions," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4814–4826, 2013.

[283] Y. Sun, Q. Wu, Y. Xu, Y. Zhang, F. Sun, and J. Wang, "Distributed channel access for device-to-device communications: A hypergraph-based learning solution," *IEEE Communications letters*, vol. 21, no. 1, pp. 180–183, 2016.

[284] B. Roberson, "The colonel blotto game," *Economic Theory*, vol. 29, no. 1, pp. 1–24, 2006.

[285] Y. Wu, B. Wang, K. R. Liu, and T. C. Clancy, "Anti-jamming games in multi-channel cognitive radio networks," *IEEE journal on selected areas in communications*, vol. 30, no. 1, pp. 4–15, 2011.

[286] Y. Wu, B. Wang, and K. R. Liu, "Optimal power allocation strategy against jamming attacks using the colonel blotto game," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pp. 1–5, IEEE, 2009.

[287] N. Namvar, W. Saad, N. Bahadori, and B. Kelley, "Jamming in the internet of things: A game-theoretic perspective," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2016.

[288] M. Labib, S. Ha, W. Saad, and J. H. Reed, "A colonel blotto game for anti-jamming in the internet of things," in *2015 IEEE global communications conference (GLOBECOM)*, pp. 1–6, IEEE, 2015.

[289] M. Guizani, A. Gouissem, K. Abualsaud, E. Yaacoub, and T. Khattab, "Combating jamming attacks in multi-channel iot networks using game theory," in *2020 3rd International Conference on Information and Computer Technologies (ICICT)*, pp. 469–474, IEEE, 2020.

[290] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Game theory for anti-jamming strategy in multichannel slow fading iot networks," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16880–16893, 2021.

[291] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Iot anti-jamming strategy using game theory and neural network," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 770–776, IEEE, 2020.

[292] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Towards secure iot networks in healthcare applications: A game theoretic anti-jamming framework," *IEEE Internet of Things Journal*, 2022.

[293] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Accelerated iot anti-jamming: A game theoretic power allocation strategy," *IEEE Transactions on Wireless Communications*, 2022.

[294] M. AbdelRaheem and M. M. Abdellatif, "Cooperative anti-jamming for secondary opportunistic networks: A colonel blotto game model," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–8, IEEE, 2017.

[295] S. Guan, J. Wang, C. Jiang, Z. Han, Y. Ren, and A. Benslimane, "Colonel blotto game aided attack-defense analysis in real-world networks," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2018.

[296] S. Guan, J. Wang, H. Yao, C. Jiang, Z. Han, and Y. Ren, "Colonel blotto games in network systems: Models, strategies, and applications," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 637–649, 2019.

[297] X. Lu, D. Xu, L. Xiao, L. Wang, and W. Zhuang, "Anti-jamming communication game for uav-aided vanets," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2017.

[298] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "Uav relay in vanets against smart jamming with reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4087–4097, 2018.

[299] W. Wang, Z. Lv, X. Lu, Y. Zhang, and L. Xiao, "Distributed reinforcement learning based framework for energy-efficient uav relay against jamming," *Intelligent and Converged Networks*, vol. 2, no. 2, pp. 150–162, 2021.

[300] L. Xiao, J. Liu, Y. Li, N. B. Mandayam, and H. V. Poor, "Prospect theoretic analysis of anti-jamming communications in cognitive radio networks," in *2014 IEEE Global Communications Conference*, pp. 746–751, IEEE, 2014.

[301] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2578–2590, 2015.

[302] L. Xiao, C. Xie, M. Min, and W. Zhuang, "User-centric view of unmanned aerial vehicle transmission against smart attacks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3420–3430, 2017.

[303] L. Xiao, D. Xu, C. Xie, N. B. Mandayam, and H. V. Poor, "Cloud storage defense against advanced persistent threats: A prospect theoretic study," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 3, pp. 534–544, 2017.

[304] A. A. A. Abass, M. Hajimirsadeghi, N. B. Mandayam, and Z. Gajic, "Evolutionary game theoretic analysis of distributed denial of service attacks in a wireless network," in *2016 Annual Conference on Information Science and Systems (CISS)*, pp. 36–41, IEEE, 2016.

[305] B. Deepak, P. S. Bharathi, and D. Kumar, "Radio frequency anti-jamming capability improvement for cognitive radio networks: An evolutionary game theoretical approach," in *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, pp. 1–6, IEEE, 2017.

[306] S. Bharathi, D. Kumar, and D. Ram, "Defence against responsive and non-responsive jamming attack in cognitive radio networks: an evolutionary game theoretical approach," *,2018,2(2018-1-29)*, vol. 2018, no. 2, 2017.

[307] Y. Bi, Y. Wu, C. Hua, and F. Zou, "Evolutionary anti-jamming game in non-orthogonal multiple access system," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.

[308] S. Bhattacharya and T. Başar, "Game-theoretic analysis of an aerial jamming attack on a uav communication network," in *proceedings of the 2010 American control conference*, pp. 818–823, IEEE, 2010.

[309] J. Parras, J. del Val, S. Zazo, J. Zazo, and S. V. Macua, "A new approach for solving anti-jamming games in stochastic scenarios as pursuit-evasion games," in *2016 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 1–5, IEEE, 2016.

[310] Y. Hu, A. Sanjab, and W. Saad, "Dynamic psychological game theory for secure internet of battlefield things (iobt) systems," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3712–3726, 2019.

[311] Y. Gao, Y. Xiao, M. Wu, M. Xiao, and J. Shao, "Game theory-based anti-jamming strategies for frequency hopping wireless communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5314–5326, 2018.

[312] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3, pp. 279–292, 1992.

[313] L. Kong, Y. Xu, Y. Zhang, X. Pei, M. Ke, X. Wang, W. Bai, and Z. Feng, "A reinforcement learning approach for dynamic spectrum anti-jamming in fading environment," in *2018 IEEE 18th International Conference on Communication Technology (ICCT)*, pp. 51–58, IEEE, 2018.

[314] S. Machuzak and S. K. Jayaweera, "Reinforcement learning based anti-jamming with wideband autonomous cognitive radios," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–5, IEEE, 2016.

[315] C. Dai, D. Xu, L. Xiao, M. Peng, and L. Sun, "Collaborative ufh-based anti-jamming broadcast with learning," in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–5, IEEE, 2017.

[316] L. Xiao, Q. Li, T. Chen, E. Cheng, and H. Dai, "Jamming games in underwater sensor networks with reinforcement learning," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2015.

[317] T. Chen, J. Liu, L. Xiao, and L. Huang, "Anti-jamming transmissions with learning in heterogenous cognitive radio networks," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 293–298, IEEE, 2015.

[318] Z. Xiao, B. Gao, S. Liu, and L. Xiao, "Learning based power control for mmwave massive mimo against jamming," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2018.

[319] G. Dubosarskii, S. Primak, and X. Wang, "Multichannel power allocation game against jammer with changing strategy," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–5, IEEE, 2018.

[320] N. Van Huynh, D. N. Nguyen, D. T. Hoang, E. Dutkiewicz, and M. Mueck, "Ambient backscatter: A novel method to defend jamming attacks for wireless networks," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 175–178, 2019.

[321] J. Peng, Z. Zhang, Q. Wu, and B. Zhang, "Anti-jamming communications in uav swarms: A reinforcement learning approach," *IEEE Access*, vol. 7, pp. 180532–180543, 2019.

[322] H. Yang, Z. Xiong, J. Zhao, D. Niyato, Q. Wu, M. Tornatore, and S. Secci, "Intelligent reflecting surface assisted anti-jamming communications based on reinforcement learning," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2020.

[323] H. Yang, Z. Xiong, J. Zhao, D. Niyato, Q. Wu, H. V. Poor, and M. Tornatore, "Intelligent reflecting surface assisted anti-jamming

communications: A fast reinforcement learning approach," *IEEE transactions on wireless communications*, vol. 20, no. 3, pp. 1963–1974, 2020.

[324] W. Li, Y. Xu, Q. Guo, Y. Zhang, X. Liu, C. Chen, and X. Song, "Joint channel selection and data scheduling in hf jamming environment: An interference-aware reinforcement learning approach," *IEEE Access*, vol. 7, pp. 157072–157084, 2019.

[325] S. Bubeck, N. Cesa-Bianchi, *et al.*, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends® in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[326] Y. Gwon, S. Dastangoo, and H. Kung, "Optimizing media access strategy for competing cognitive radio networks," in *2013 IEEE Global Communications Conference (GLOBECOM)*, pp. 1215–1220, IEEE, 2013.

[327] S. Dastangoo, C. E. Fossa, Y. L. Gwon, and H.-T. Kung, "Competing cognitive resilient networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 1, pp. 95–109, 2016.

[328] Q. Wang, K. Ren, and P. Ning, "Anti-jamming communication in cognitive radio networks with unknown channel statistics," in *2011 19th IEEE International Conference on Network Protocols*, pp. 393–402, IEEE, 2011.

[329] H. Su, Q. Wang, K. Ren, and K. Xing, "Jamming-resilient dynamic spectrum access for cognitive radio networks," in *2011 IEEE International Conference on Communications (ICC)*, pp. 1–5, IEEE, 2011.

[330] Q. Wang, K. Ren, P. Ning, and S. Hu, "Jamming-resistant multi-radio multichannel opportunistic spectrum access in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8331–8344, 2015.

[331] P. Zhou, Q. Yan, K. Wang, Z. Xu, S. Ji, and K. Bian, "Jamsa: A utility optimal contextual online learning framework for anti-jamming wireless scheduling under reactive jamming attack," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1862–1878, 2019.

[332] Q. Wang, P. Xu, K. Ren, and X.-Y. Li, "Towards optimal adaptive ufh-based anti-jamming wireless communication," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 16–30, 2011.

[333] P. Zhou, Q. Wang, W. Wang, Y. Hu, and D. Wu, "Near-optimal and practical jamming-resistant energy-efficient cognitive radio communications," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2807–2822, 2017.

[334] P. Zhou, J. Xu, W. Wang, Y. Hu, D. O. Wu, and S. Ji, "Toward optimal adaptive online shortest path routing with acceleration under jamming attack," *IEEE/ACM transactions on networking*, vol. 27, no. 5, pp. 1815–1829, 2019.

[335] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[336] S. Liu, Y. Xu, X. Chen, X. Wang, M. Wang, W. Li, Y. Li, and Y. Xu, "Pattern-aware intelligent anti-jamming communication: A sequential deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 169204–169216, 2019.

[337] H. Han, X. Wang, F. Gu, W. Li, Y. Cai, Y. Xu, and Y. Xu, "Better late than never: Gan-enhanced dynamic anti-jamming spectrum access with incomplete sensing information," *IEEE Wireless Communications Letters*, vol. 10, no. 8, pp. 1800–1804, 2021.

[338] H. Han, Y. Xu, Z. Jin, W. Li, X. Chen, G. Fang, and Y. Xu, "Primary-user-friendly dynamic spectrum anti-jamming access: A gan-enhanced deep reinforcement learning approach," *IEEE Wireless Communications Letters*, vol. 11, no. 2, pp. 258–262, 2021.

[339] X. Liu, Y. Xu, Y. Cheng, Y. Li, L. Zhao, and X. Zhang, "A heterogeneous information fusion deep reinforcement learning for intelligent frequency selection of hf communication," *China communications*, vol. 15, no. 9, pp. 73–84, 2018.

[340] M. A. Aref and S. K. Jayaweera, "Robust deep reinforcement learning for interference avoidance in wideband spectrum," in *2019 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)*, pp. 1–5, IEEE, 2019.

[341] Y. Li, X. Wang, D. Liu, Q. Guo, X. Liu, J. Zhang, and Y. Xu, "On the performance of deep reinforcement learning-based anti-jamming method confronting intelligent jammer," *Applied Sciences*, vol. 9, no. 7, p. 1361, 2019.

[342] J. Xu, H. Lou, W. Zhang, and G. Sang, "An intelligent anti-jamming scheme for cognitive radio based on deep reinforcement learning," *IEEE Access*, vol. 8, pp. 202563–202572, 2020.

[343] X. Chang, Y. Li, Y. Zhao, Y. Du, and D. Liu, "An improved anti-jamming method based on deep reinforcement learning and feature engineering," *IEEE Access*, vol. 10, pp. 69992–70000, 2022.

[344] Y. Bi, Y. Wu, and C. Hua, "Deep reinforcement learning based multi-user anti-jamming strategy," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2019.

[345] X. Lu, L. Xiao, C. Dai, and H. Dai, "Uav-aided cellular communications with deep reinforcement learning against jamming," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 48–53, 2020.

[346] Y. Chen, Y. Li, D. Xu, and L. Xiao, "Dqn-based power control for iot transmission against jamming," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, IEEE, 2018.

[347] N. Ma, K. Xu, X. Xia, C. Wei, Q. Su, M. Shen, and W. Xie, "Reinforcement learning-based dynamic anti-jamming power control in uav networks: An effective jamming signal strength based approach," *IEEE Communications Letters*, vol. 26, no. 10, pp. 2355–2359, 2022.

[348] W. Li, Y. Qin, Z. Feng, H. Han, J. Chen, and Y. Xu, ""advancing secretly by an unknown path": A reinforcement learning-based hidden strategy for combating intelligent reactive jammer," *IEEE Wireless Communications Letters*, 2022.

[349] L. Xiao, D. Jiang, Y. Chen, W. Su, and Y. Tang, "Reinforcement-learning-based relay mobility and power allocation for underwater sensor networks against jamming," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 3, pp. 1148–1156, 2019.

[350] G. Han, L. Xiao, and 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)*, pp. 2087–2091, IEEE, 2017.

[351] W. Li, J. Wang, L. Li, Q. Peng, W. Huang, X. Chen, and S. Li, "Secure and reliable downlink transmission for energy-efficient user-centric ultra-dense networks: An accelerated drl approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 8978–8992, 2021.

[352] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, ""jam me if you can:" defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communications," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2603–2620, 2019.

[353] P. D. Thanh, H. T. H. Giang, and I.-P. Hong, "Anti-jamming ris communications using dqn-based algorithm," *IEEE Access*, vol. 10, pp. 28422–28433, 2022.

[354] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Defeating reactive jammers with deep dueling-based deception mechanism," in *ICC 2021-IEEE International Conference on Communications*, pp. 1–6, IEEE, 2021.

[355] N. Abuzainab, T. Erpek, K. Davaslioglu, Y. E. Sagduyu, Y. Shi, S. J. Mackey, M. Patel, F. Panettieri, M. A. Qureshi, V. Isler, *et al.*, "Qos and jamming-aware wireless networking using deep reinforcement learning," in *MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM)*, pp. 610–615, IEEE, 2019.

[356] Z. Li, Y. Lu, X. Li, Z. Wang, W. Qiao, and Y. Liu, "Uav networks against multiple maneuvering smart jamming with knowledge-based reinforcement learning," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12289–12310, 2021.

[357] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[358] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[359] G. Chen, Y. Zhan, Y. Chen, L. Xiao, Y. Wang, and N. An, "Reinforcement learning based power control for in-body sensors in wbans against jamming," *IEEE Access*, vol. 6, pp. 37403–37412, 2018.

[360] C. Han and Y. Niu, "Multi-regional anti-jamming communication scheme based on transfer learning and q learning," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 13, no. 7, pp. 3333–3350, 2019.

[361] H. T. Thien, V.-H. Vu, and I. Koo, "A transfer games actor–critic learning framework for anti-jamming in multi-channel cognitive radio networks," *IEEE Access*, vol. 9, pp. 47887–47900, 2021.

[362] L. Xiao, D. Jiang, D. Xu, H. Zhu, Y. Zhang, and H. V. Poor, "Two-dimensional antijamming mobile communication based on reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9499–9512, 2018.

[363] L. Xiao, Y. Ding, J. Huang, S. Liu, Y. Tang, and H. Dai, "Uav anti-jamming video transmissions with qoe guarantee: A reinforcement learning-based approach," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5933–5947, 2021.

[364] X. Lu, L. Xiao, G. Niu, X. Ji, and Q. Wang, "Safe exploration in wireless security: A safe reinforcement learning algorithm with hierarchical structure," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 732–743, 2022.
[365] Q. Wang and M. Liu, "Learning in hide-and-seek," *IEEE/ACM transactions on networking*, vol. 24, no. 2, pp. 1279–1292, 2015.
[366] J. Dams, M. Hoefer, and T. Kesselheim, "Jamming-resistant learning in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2809–2818, 2015.
[367] R. Basomingera and Y.-J. Choi, "Deep multi-task conditional and sequential learning for anti-jamming," *IEEE Access*, vol. 9, pp. 123194–123207, 2021.
[368] L. Zhang, F. Restuccia, T. Melodia, and S. M. Pudlewski, "Taming cross-layer attacks in wireless networks: a bayesian learning approach," *IEEE Transactions on Mobile Computing*, vol. 18, no. 7, pp. 1688–1702, 2018.
[369] N. I. Mowla, N. H. Tran, I. Doh, and K. Chae, "Afrl: Adaptive federated reinforcement learning for intelligent jamming defense in fanet," *Journal of Communications and Networks*, vol. 22, no. 3, pp. 244–258, 2020.