IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 December 2022, accepted 9 December 2022, date of publication 15 December 2022,
date of current version 22 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3230004

RESEARCH ARTICLE

A Multi-Domain Anti-Jamming Scheme Based on Bayesian Stackelberg Game With Imperfect Information

YONGCHENG LI[1], KANGZE LI[2], ZHENZHEN GAO[2,3], AND CHUNLEI ZHENG[4]

[1]State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE), Luoyang 471003, China
[2]School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[3]National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
[4]Shanghai Institute of Microsystem and Information Technology Chinese Academy of Sciences, Shanghai 200050, China

Corresponding author: Zhenzhen Gao (zhenzhengao@xjtu.edu.cn)

This work was supported in part by the Open Research Found of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE) under Grant 2022K0202A; in part by the National Natural Science Foundation of China under Grant 62071367; and in part by the Open Research Found of National Mobile Communications Research Laboratory, Southeast University, under Grant 2020D11.

ABSTRACT To deal with the smart jammer which can sense the legitimate communication and adjust its jamming policy, an anti-jamming game scheme is proposed in this paper to jointly optimize the transmission power and frequency hopping period. Considering that both parties in the game can only get the other's channel gains in probabilistic form, the interaction between the legitimate transmitter and the jammer is modeled as Bayesian Stackelberg game, where the legitimate transmitter is set as the leader while the jammer acts as the follower. The transmitter and the jammer determine their optimal strategies in power domain and time domain to maximize their own utilities which are formulated based on the spectrum efficiency. Besides, the imperfect information including observation errors and the bounded rationality of the jammer is considered when formulating the utility functions. By using backward induction, the optimal solutions in time-power domain are obtained for the transmitter and the jammer. Simulation results show that the proposed multi-domain game scheme outperforms the single-domain game schemes and the multi-domain random scheme. Moreover, the impact of imperfect information is discussed through simulations.

INDEX TERMS Anti-jamming, multi-domain, Stackelberg game, imperfect information.

I. INTRODUCTION

Due to the open nature of the wireless channel, the wireless communication between the legitimate transmitter-receiver pair is exposed to the threat of jamming, which seriously affects the communication quality [1]. To deal with jamming attacks, many anti-jamming techniques in frequency domain [2] and power domain [3], [4], [5] have been proposed. The main anti-jamming technology in the frequency domain is spread spectrum technology, such as frequency hopping, direct sequence spread spectrum, and adaptive spread spectrum. Another widely used anti-jamming technology is applied in the power domain, which is mainly realized by changing the transmit power of the transmitter and the jammer.

Being equipped with spectrum sensing technology, a smart jammer can actively sense the legitimate communication and adjust its jamming policy [6]. The smart jammer causes a great challenge to the current anti-jamming technologies. Since the smart jammer can adapt itself to its sensing results including the transmission frequency band and the transmission power, it is important for the legitimate transmitter to adjust the transmit frequency band and the transmit power nimbly. Therefore, to effectively deal with the smart jammer, there are two fundamental problems to be solved: How to model the competition between the jammer and the transmitter? How to adaptively adjust the transmission frequency band and the transmission power?

The associate editor coordinating the review of this manuscript and approving it for publication was Chan Hwang See.

For the first problem, game theory is an useful tool to model the competition between two players [7]. For the second problem, it is intuitive that an optimal value exists respectively for the frequency hopping period and the transmission power. Specifically, if the frequency hopping period is too long, it will be easier for the jammer to detect the communication signal and perform effective jamming. If the frequency hopping period is too short, the effective communication time will be decreased. Similarly, larger transmission power brings better communication quality, but it is also easier to be detected by the jammer. Therefore, it is necessary to jointly find the optimal values in time domain and power domain for the frequency hopping period and the transmit power.

A. RELATED WORK

To deal with jamming attacks, many related techniques have been proposed [6], [8], [9], [10]. When the competition between the transmitter and the jammer is considered, game theory [11] is a powerful mathematical tool to analyze the interaction, especially the Stackelberg game model, which is usually used to model the hierarchical competition between the transmitter and the jammer.

A power-domain anti-jamming scheme has been designed based on Stackelberg game in [4] for cooperative anti-jamming communications. In [9], a power control Stackelberg game was proposed to cope with a smart jammer in wireless communication systems. Furthermore, the authors in [6] studied the use of power control methods to resist intelligent jamming in cognitive radio networks with observation errors, and derived the Stackelberg Equilibrium between the users and jammers. In [3], an anti-jamming Bayesian Stackelberg game with incomplete information was proposed, and the optimal transmission power based on duality optimization theory was derived. Stackelberg game was used in [12] to solve the anti-jamming problem in the UAV communication network where the drones interfere with each other.

The above work mainly considered the power-domain anti-jamming technologies based on game theory. In addition to the power domain, there are also anti-jamming techniques for wireless communications in other domains, such as frequency domain and spatial domain. The authors considered jamming and anti-jamming in interference channels by way of smart hopping and obtained smart channel hopping sequences for both the jammer and the target-transmitter in [13]. In [14], the authors investigated the problem of dynamic spectrum access for canonical wireless networks with time-varying channels. They formulated the interactions among the users in the time-varying environment as a non-cooperative game. A bimatrix game framework was developed in [15] to model the interaction process between the transmitter and the jammer in frequency hopping wireless communications, where each player made its decision on whether to stay on the current channel or hop to a new one. Authors in [16] adopted the zero-sum game and a deep Q-network algorithm to solve the anti-jamming issue in the frequency-spatial domain in cognitive networks. A multi-domain anti-jamming scheme was proposed in [17], where a Stackelberg power game was formulated in the power domain to fight against the jamming attacks, and a multi-armed bandit-based channel selection with a channel switching cost and unknown channel availability state information was formulated in the spectrum domain. The shortcomings of the separate application of FH and transmission rate adaptation methods were discussed in [18], and the idea of joint use of the two technologies was proposed to prevent interference. It was proved that multi-domain anti-jamming technology has better performance and greater flexibility than single-domain anti-jamming technology [17], [18]. Authors in [19] proposed a multi-domain anti-jamming strategy using Stackelberg game for wireless relay networks with perfect information and demonstrated that the proposed multi-domain strategy outperformed single-domain schemes.

B. CONTRIBUTIONS

In this paper, a multi-domain anti-jamming game problem is solved by jointly optimizing the time-domain and power-domain strategies. Since there exist observation errors in practical applications, the jammer may deviate from the optimal strategy but choose sub-optimal strategies, which make the jammer act with bounded rationality. Besides, it is difficult for the transmitter and the jammer to obtain perfect channel information about the other part in the game, imperfect channel information is also considered. The imperfect channel information, the observation errors and the bounded rationality make the obtained information of the game model become imperfect. It is necessary to model the imperfect information and analyze the influence of the imperfect information in the considered multi-domain anti-jamming game.

The contributions of this paper can be summarized as follows:

- The power domain and the time domain parameters are jointly optimized in the proposed multi-domain anti-jamming game scheme. As we have analyzed that, besides the widely considered power domain, the frequency hopping period in time domain is also critical to the communication and jamming performance. Thus both the time domain and the power domain parameters are considered when establishing the utility functions of the anti-jamming game.
- The imperfect information is considered in the proposed multi-domain anti-jamming game scheme. Bayesian Stackelberg game is used to model the competition of the transmitter and the jammer when imperfect channel information is considered. The observation errors and the bounded rationality are modeled and analyzed. Some discussions and insights are given to reflect the influence of the imperfect information.

C. ORGANIZATION AND NOTATIONS

The remaining of this paper is organized as follows. We present the Bayesian Stackelberg game model and give the problem formulation in Section II. The proposed multi-domain anti-jamming game is solved in Section III. Simulation results and discussions are given in Section IV. Finally, Section V concludes this paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

We consider that the wireless communication between a transmitter S and a receiver D is attacked by a smart jammer J. Let h_{sd}, h_{sj}, h_{jd} denote the channel gains between S and D, S and J, J and D, respectively. P_s and P_j denote the transmit power of the transmitter and the jammer. Considering the maximum power constraint, let P_{smax} and P_{jmax} represent the maximum transmit power of the transmitter and the jammer, respectively. T represents the frequency hopping period of the wireless frequency hopping communication system.

In this paper, we consider a frequency hopping communication system as shown in Fig. 1. Let F denote the frequency set that the transmitter can use, $F = [f_1, f_2, f_3, \ldots, f_M]$. The frequency hopping signal hops pseudo-randomly in M adjacent sub-bands. The frequency hopping period satisfies $T \in [0, T_{max}]$, and T_{max} is the maximum frequency hopping period. The transmitter can adaptively adjust the frequency hopping period within the range of $[0, T_{max}]$ according to the parameters of the smart jammer. Due to the limitations of devices, there exists an unstable transient process when the transmitter switches from one frequency band to another, and the duration of this process is related to the hardware [20]. Denote the duration of the transient process as frequency switching time T_0, the transmitter and the receiver remain silent during T_0.

To perform effective jamming, the smart jammer actively senses the transmission strategy of the transmitter. The time required for this sensing process is defined as signal detection time, which is represented by T_E. In this process, the smart jammer adaptively adjusts its detection time and jamming power according to its sensing results. Similarly, the transmitter will adjust its transmission power and frequency hopping period when it finds itself under jamming attack. In each frequency hopping period, the smart jammer performs jamming attack as soon as the legitimate transmission is detected, otherwise the jammer will remain silient.Therefore, each frequency hopping period can be divided into two parts: the signal detection time T_E and the jamming time T_I, and $T = T_E + T_I$. The signal detection probability is related to the detection time T_E and the transmit signal power P_s. Assume that the energy detection is used at the jammer, the signal detection probability of the transmitter's frequency hopping signal can be written as [21]

$$p_d = \sum_{m=0}^{M-1} (-1)^m \binom{M-1}{m} \frac{1}{m+1} \exp\left(\frac{-m}{4(m+1)} \mathrm{SNR}_j T_E\right) \approx 1 - \frac{M-1}{M} \exp\left(-\frac{1}{4} \mathrm{SNR}_j T_E\right), \tag{1}$$

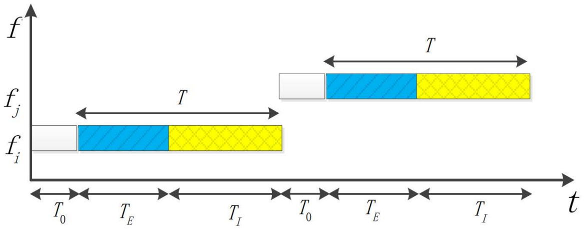


FIGURE 1. The structure of the frequency hopping signal.

where $\binom{M-1}{m}$ is the number of m-combination of a set with $M-1$ elements, SNR_j is the received signal-to-noise-ratio (SNR) at J, which can be expressed as

$$\mathrm{SNR}_j = \frac{h_{sj} P_s}{\delta_J^2}, \tag{2}$$

where δ_J^2 is the noise power of the additive white Gaussian noise (AWGN) at J. We can find out that higher detection probability can be obtained by the smart jammer if the detection time or the transmit signal power increases.

When the smart jammer does not detect any communication signal, it will keep silent, and the received SNR at D can be expressed as

$$\mathrm{SNR}_d = \frac{h_{sd} P_s}{\delta_D^2}, \tag{3}$$

where δ_D^2 is the noise power of the AWGN at D. When the smart jammer detects the communication signal, it will perform jamming and the received signal-to-interference-plus-noise ratio (SINR) at D can be written as

$$\mathrm{SINR}_d = \frac{h_{sd} P_s}{\delta_D^2 + h_{jd} P_j}. \tag{4}$$

Considering that in practical applications, the receiver and the jammer are usually in a similar environment, we can set $\delta_J^2 = \delta_D^2 = \delta^2$ without loss of generality.

B. GAME MODEL

1) IMPERFECT CHANNEL INFORMATION

Here we use the path loss model [22] to model the channel gains. It is difficult for the transmitter to get the perfect channel gains between the jammer and the receiver in practice. Therefore, we assume that the transmitter only gets the probability distribution of the channel gains of the jammer as in [3] and [5]. The receiver can observe the behavior of the jammer and obtain the probability distribution of the jammer's location from the historical information. This probability distribution information can be feedback to the transmitter. Based on the path loss model, the transmitter can obtain the imperfect channel gains as in Assumption 1.

Assumption 1: For the transmitter, since the channel gain is modeled by the channel path loss, the channel gains of the $S-J$ link and $J-D$ link are related to the positions of J. Assume that the jammer has N possible locations, h_{sj} and h_{jd} have N possible states denoted by $h_{sj}(n)$ and $h_{jd}(n)$ ($n = 1, 2, \ldots, N$) respectively. The probability for $h_{sj}(n)$ or $h_{jd}(n)$ is β_n and $\sum_{n=1}^{N} \beta_n = 1$.

Similarly, the jammer can not get the perfect channel gain between S and D. But the probability distribution of the channel gains is available at the jammer as in [3] and [5].

Assumption 2: For the smart jammer, the channel gain h_{sd} has Q possible states which are $h_{sd}(q)$, $(q = 1, 2, \ldots Q)$. The probability of $h_{sd}(q)$ is η_q, and $\sum_{q=1}^{Q} \eta_q = 1$.

2) OBSERVATION ERROR AND BOUNDED RATIONALITY

During the game, the follower will observe the leader's transmit power P_s and the frequency hopping period T. Inspired by [23], the observation error factors for P_s and T can be expressed as $\varepsilon_1 = \left|\widetilde{P}_s - P_s\right| / P_s$, $\varepsilon_2 = \left|\widetilde{T} - T\right| / T$, respectively, where $\widetilde{P}_s$ and $\widetilde{T}$ represent the observations of the jammer for the transmitter's transmit power and frequency hopping period. ε_1 and ε_2 are used to describe the inaccurate observations, and $\varepsilon_i \in [0, 1]$, $i = 1, 2$. If $\varepsilon_i = 0$, it means that the jammer can perfectly observe the transmitter's strategy.

Due to the bounded rationality of the jammer, it may not strictly choose the optimal strategy. The jammer's bounded rationality will influence the transmitter's decision. In [23] and [24], the bounded rationality has been modeled for discrete strategies. In our paper, continuous strategies in time domain and power domain are considered. Take the time domain as an example. Let T_E^* denote the optimal strategy that maximize the jammer's utility. Due to the reasons like inaccurate observations, the jammer may deviate from T_E^* and behaves with bounded rationality. Due to the bounded rationality, the suboptimal strategy adopted by the jammer can be written as $\widehat{T}_E = T_E^* + T_{E0}$, where T_{E0} illustrates the degree of deviation, and we use a zero-mean Gaussian random variable to model this deviation. Smaller variance of T_{E0} represents higher rationality degree. If the variance is zero, it means that the jammer is perfectly rational. The subsequent simulations show that the proposed modeling of the bounded rationality is consistent with the conclusions obtained in [25].

3) BAYESIAN STACKELBERG GAME

In this section, the competition between S and J will be modeled by using Bayesian Stackelberg game, where S acts as the leader and J is the follower. Both S and J try to find the optimal parameters in time domain and power domain to maximize their own utilities. Mathematically, we denote the strategy space of the transmitter and the jammer in time domain and power domain as $\mathcal{S}$ and $\mathcal{J}$ respectively, specifically, $\mathcal{S} = \{T, P_s\}$ and $\mathcal{J} = \{T_E, P_j\}$. Let μ_s and μ_j represent the utility functions of the transmitter and the jammer, respectively.

A utility function which can intuitively and accurately reflect the communication performance is preferred. Therefore, the spectrum efficiency of the communication system is used as the performance metric when formulating the utility function. For S, it tries to maximize the transmission spectrum efficiency, while J tries to degrade the communication and minimize the communication spectrum efficiency. Considering the uncertainties of jamming channel gains h_{sj}, h_{jd} and the bounded rationality, the utility function of the transmitter can be formulated based on Bayesian Stackelberg game as $\mu_s = \sum_{n=1}^{N} \beta_n \mu_{s|n}$, where $\mu_{s|n}$ is the conditional utility given the jamming channel gain $h_{sj}(n)$, $h_{jd}(n)$, $n = 1, 2, \cdots, N$. The conditional utility can be expressed as

$$\begin{aligned}\mu_{s|n} = {} & \frac{T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}P_s}{\delta^2}) \\ & +(1 - \widehat{p}_{d|n}) \times \frac{T-T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}P_s}{\delta^2}) \\ & +\widehat{p}_{d|n} \times \frac{T-T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}P_s}{\delta^2 + h_{jd}(n)P_j}), \end{aligned} \quad (5)$$

where $\widehat{p}_{d|n}$ is the estimated detection probability at S when then jamming channel uncertainty and the bounded rationality are taken into consideration, the estimated detection probability at S is

$$\widehat{p}_{d|n} = 1 - \frac{M-1}{M} \exp\left(-\frac{1}{4}\frac{h_{sj}(n)P_s}{\delta^2}\widehat{T}_E\right). \quad (6)$$

The conditional utility function in Equ. (5) consists of three terms. The first term is the effective spectrum efficiency when the jammer is performing signal detection during the detection time T_E. The second term represents the effective spectrum efficiency when the jammer detects no signal on the considered frequency band during the detection time T_E. In the above two cases the jammer keeps silent and the receiver will not be interfered by the jammer. The third term represents the case that the jammer detects the communication signal successfully and carries out jamming with jamming power P_j.

As a leader, the transmitter can determine the optimal transmit power P_s^* and the optimal frequency hopping period T^* from the following optimization problem:

$$(P_s^*, T^*) = \arg \max_{0 \leqslant P_s \leqslant P_{smax}, 0 \leqslant T \leqslant T_{max}} \mu_s(P_s, T). \quad (7)$$

As for the jammer, considering the channle uncertainty of h_{sd} and the observation errors, the jammer's utility can be written as $\mu_j = \sum_{q=1}^{Q} \eta_q \mu_{j|q}$, where $\mu_{j|q}$ is the conditional utility given the communication channel gain $h_{sd}(q)$, $q = 1, 2, \cdots, Q$. The conditional utility at J is as follows

$$\begin{aligned}\mu_{j|q} = -\Big(& \frac{T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}(q)P_s}{\delta^2}) \\ & +(1 - \widetilde{p}_d) \times \frac{T-T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}(q)P_s}{\delta^2}) \\ & +\widetilde{p}_d \times \frac{T-T_E}{T+T_0} \times \ln(1 + \frac{h_{sd}(q)P_s}{\delta^2 + h_{jd}P_j})\Big), \end{aligned} \quad (8)$$

where $\widetilde{p}_d$ is the detection probability at J when inaccurate observations are considered, the detection probability at J can be written as

$$\widetilde{p}_d = 1 - \frac{M-1}{M} \exp\left(-\frac{1}{4}\frac{h_{sj}\widetilde{P}_s}{\delta^2} T_E\right). \quad (9)$$

Assume that the jammer has sufficient power supply, the jamming power cost is not considered in the jammer's utility.

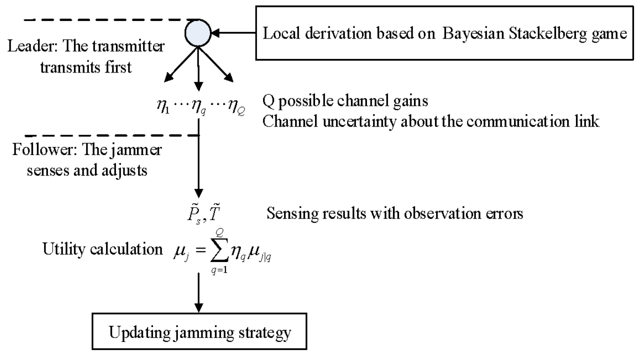

Leader: The transmitter transmits first
Local derivation based on Bayesian Stackelberg game
$\eta_1 \cdots \eta_q \cdots \eta_Q$
Q possible channel gains
Channel uncertainty about the communication link
Follower: The jammer senses and adjusts
$\tilde{P}_s, \tilde{T}$
Sensing results with observation errors
Utility calculation $\mu_j = \sum_{q=1}^{Q} \eta_q \mu_{j|q}$
Updating jamming strategy

FIGURE 2. Major processes of both parties in Bayesian Stackelberg game.

Then the jammer will attack the transmitter with the maximum power of P_{jmax}. For the jammer, the parameter to be optimized is the signal detection time T_E. As a follower, the jammer can obtain the optimal signal detection time $T_E{}^*$ from the following optimization problem

$$T_E^* = \arg \max_{0 \leqslant T_E < T} \mu_j(T_E). \tag{10}$$

The major processes at S and J are shown in Fig. 2. The local derivation at the transmitter in Fig. 2 includes the following steps: (1) the transmitter predicts the jammer's strategy locally based on the game model and some parameters of the jammer; (2) the transmitter derives its optimal strategy based on the predicted jamming strategy to maximize its utility; (3) the transmitter transmits according to its optimal strategy. As a follower, the jammer performs sensing to estimate the transmit power and the frequency hopping period of the transmitter. Based on these estimations and the common knowledge of the utilities, the jammer will calculate its utility and find the optimal jamming strategy to maximize its utility. Finally, the jammer adjusts its attack strategy according to its solution. The detailed derivation is given in the following section.

III. THE MULTI-DOMAIN OPTIMIZATION STRATEGY BASED ON BAYESIAN STACKELBERG GAME

In a Bayesian Stackelberg game, backward induction is an effective method to obtain the optimal solution [26], [27]. In this section, we first solve the follower sub-game through mathematical derivation, and then find the optimal strategy for the leader.

A. FOLLOWER SUB-GAME

In this part, we first solve the follower sub-game, which is to find the optimal signal detection time T_E^*. For the imperfect channel information case, the jammer's utility function μ_j is simplified as follows

$$\mu_j = -\sum_{q=1}^{Q} \eta_q \Big(\frac{T_E}{T+T_0} \times B_q + (1-p_d) \times \frac{T-T_E}{T+T_0} \times B_q + p_d \times \frac{T-T_E}{T+T_0} \times A_q \Big), \tag{11}$$

where

$$A_q = \ln(1 + \frac{h_{sd}(q)P_s}{\delta^2 + h_{jd}P_j}),\ B_q = \ln(1 + \frac{h_{sd}(q)P_s}{\delta^2}). \tag{12}$$

Here we can see that $A_q < B_q$. For the signal detection probability, the first-order Taylor expansion of the exponential function is performed to obtain

$$p_d = \frac{1}{M} + \frac{(M-1)h_{sj}P_s}{4M\delta^2} \times T_E. \tag{13}$$

Bring A, B and p_d into μ_j, we can get

$$\mu_j = \sum_{q=1}^{Q} \eta_q \Big[\frac{(A_q - B_q)(M-1)h_{sj}P_s}{4M\delta^2(T+T_0)} \times T_E{}^2 - (A_q - B_q)\Big(\frac{(M-1)h_{sj}P_sT}{4M\delta^2} - \frac{1}{M(T+T_0)}\Big) \times T_E - \frac{(A_q - B_q)T}{M(T+T_0)} - \frac{T}{(T+T_0)} \times B_q \Big]. \tag{14}$$

By observing the above formula, it can be found that μ_j is a quadratic function of T_E. Because $A_q < B_q$, we can know that $\frac{(A_q - B_q)(M-1)h_{sj}P_s}{4M\delta^2(T+T_0)}$ is less than 0. According to the properties of the quadratic function, it can be known that μ_j has a unique maximum value. μ_j takes the maximum value if and only if T_E takes the value of the axis of symmetry. According to the expression of the symmetry axis of the quadratic function, we can get

$$T_E^* = \frac{T}{2} - \frac{2\delta^2}{P_s h_{sj}(M-1)}. \tag{15}$$

It can be seen from Equ. (15) that the optimal detection time T_E^* we found does not include A_q and B_q. It means that the uncertainty of h_{sd} does not affect the optimal detection time T_E^* for the jammer. When the observation errors are considered, the signal detection time of the jammer becomes

$$\tilde{T}_E^* = \frac{\tilde{T}}{2} - \frac{2\delta^2}{\tilde{P}_s h_{sj}(M-1)}. \tag{16}$$

B. LEADER SUB-GAME

In this part, we try to optimize the transmitter's strategy, that is, to find the transmitter's optimal transmit power P_s^* and the optimal frequency hopping period T^*. We first consider the ideal case that the jammer is fully rational. By substituting T_E with T_E^* in μ_s, the transmitter's utility function based on Bayesian Stackelberg game becomes

$$\mu_s = \sum_{n=1}^{N} \beta_n \Big[\frac{(A_n - B)(M-1)h_{sj}(n)P_s}{16M\delta^2} \times \frac{T^2}{T+T_0} + (\frac{A_n - B}{2M} + B) \times \frac{T}{T+T_0} + \frac{(A_n - B)\delta^2}{M(M-1)h_{sj}(n)P_s} \times \frac{1}{T+T_0} \Big], \tag{17}$$

where

$$A_n = \ln(1 + \frac{h_{sd}P_s}{\delta^2 + h_{jd}(n)P_j}),\quad B = \ln(1 + \frac{h_{sd}P_s}{\delta^2}). \tag{18}$$

Here, we first find the optimal frequency hopping period T^*. Taking the partial derivative of μ_s with respect to T, we can get:

$$\begin{aligned}\frac{\partial\mu_s}{\partial T} &= \sum_{n=1}^{N}\beta_n\Big[\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}\times\frac{T^2}{(T+T_0)^2}\\ &\quad +2T_0\times\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}\times\frac{T}{(T+T_0)^2}\\ &\quad +((\frac{A_n-B}{2M}+B)\times T_0-\frac{(A_n-B)\delta^2}{M(M-1)h_{sj}(n)P_s})\\ &\quad \times\frac{1}{(T+T_0)^2}\Big]. \end{aligned} \tag{19}$$

Observing the expression for the first-order partial derivative obtained above, we can see that its graph is similar to that of a quadratic function with an opening downward. The image has two intersections with the X-axis, so the first partial derivative has two zeros. We can determine that as T increases from zero, the trend of μ_s is to decrease first, then increase and then decrease. Thus we can know that the maximum value of μ_s is obtained at the right-hand zero point of the first-order partial derivative. Let $\frac{\partial\mu_s}{\partial T}=0$ and divide both sides of the equation by $\frac{1}{(T+T_0)^2}$ to get the following equation:

$$\begin{aligned}&\sum_{n=1}^{N}\beta_n\Big[\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}\times T^2\\ &\quad +2T_0\times\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}\times T\\ &\quad +(\frac{A_n-B}{2M}+B)\times T_0-\frac{(A_n-B)\delta^2}{M(M-1)h_{sj}(n)P_s}\Big]=0. \end{aligned} \tag{20}$$

The equation obtained above is a typical one-variable quadratic equation, so we can find the expressions for the two solutions through the root-finding formula of the one-variable quadratic equation. Given the jamming channel gains, we respectively denote the coefficients of the quadratic term, the coefficient of the first-order term and the constant term in the above formula as a_n, b_n and c_n as follows,

$$a_n=\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}, \tag{21}$$

$$b_n=2T_0\times\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}, \tag{22}$$

$$c_n=(\frac{A_n-B}{2M}+B)\times T_0-\frac{(A_n-B)\delta^2}{M(M-1)h_{sj}(n)P_s}. \tag{23}$$

Through the formula for finding the root of the quadratic equation, we can get

$$T=\sum_{n=1}^{N}\beta_n(\frac{-b_n\pm\sqrt{b_n^2-4a_nc_n}}{2a_n}). \tag{24}$$

Since $a<0$, T^* needs to take the zero point on the right. That is, take the larger one of the two solutions, so

$$T^*=\sum_{n=1}^{N}\beta_n(\frac{-b_n-\sqrt{b_n^2-4a_nc_n}}{2a_n}). \tag{25}$$

It will be found from the simulation results that the optimal time domain solutions $T_E{}^*$ in Equ. (15) and T^* in Equ. (25) are always positive for considered simulation settings.

When the bounded rationality of the jammer is considered, the jammer may adopt the signal detection time $\widehat{T_E^*}=T_E^*+T_{E0}$ instead of T_E^*, and the estimated detection probability at S can be calculated by using Equ.(6). In this case, the transmitter's utility function becomes

$$\begin{aligned}\mu_s' &= \sum_{n=1}^{N}\beta_n\Big[\frac{(A_n-B)(M-1)h_{sj}(n)P_s}{16M\delta^2}\times\frac{T^2}{T+T_0}\\ &\quad +(\frac{A_n-B}{2M}+B)\times\frac{T}{T+T_0}\\ &\quad +\frac{(A_n-B)\delta^2}{M(M-1)h_{sj}(n)P_s}\times\frac{1}{T+T_0}\\ &\quad +\frac{(A_n-B)}{T+T_0}(\frac{(M-1)h_{sj}(n)P_sT}{8M\delta^2}+\frac{1}{2M})\times T_{E0}\Big]. \end{aligned} \tag{26}$$

In this case, the third term of μ_s' contains a zero-mean Gaussian random variable, which makes it quite difficult to derive a closed-form solution for T and P_s. However, after taking expectation with respect to T_E, we can find that the average value of the optimal frequency hopping period for this case is the solution given in Equ. (25).

Substituting T^* into Equ. (17) or Equ. (26), we get an expression for μ_s with respect to P_s when the jammer is perfect rational or bounded rational. Because the expression of μ_s is too complicated, we cannot obtain the analytical expression of the transmitter's optimal transmit power $P_s{}^*$ through mathematical derivation. There are many algorithms to find the optimal solution, such as genetic algorithm, ant colony algorithm, annealing algorithm, etc. Since this paper focuses on the modeling of the imperfect information and Bayesian Stackelberg game, genetic algorithm (GA) is used directly to solve this optimization problem. A brief description about the steps of GA is described as follows.

1) INITIAL POPULATION

First, we randomly create an initial population $pop(0)$ in the solution space, which includes n individuals, and each individual is represented by $S_i^0=P_{s_i^0}, (i=1,2,\ldots,n)$. For each individual S_i^0, the corresponding utility value of the transmitter can be calculated by Equ. (17) or Equ. (26).

2) EVALUATION

Second, we establish a reasonable fitness calculation function and calculate the fitness corresponding to each individual. For the calculation of fitness, we use the following formula:

$$R(S_i^k)=\frac{1}{f_{xMMax}-f_x(S_i^k)}, \tag{27}$$

where S_i^k is the individual of kth generation, $R(S_i^k)$ is the fitness of each individual S_i^k, $f_x(\cdot)$ is the utility function and f_{xMMax} represents any value greater than the maximum value of the function. The closer its value is to the maximum value of the function, the better the convergence of the algorithm will be.

3) SELECT

Based on the fitness of individuals in the population, individuals with higher fitness value are selected for direct inheritance, or to generate new individuals of the next generation through pairing and crossover. Here we use the roulette method to construct the selection operator.

We divide individual fitness by the sum of all individual fitness to get the probability $p(S_i^k)$ of individual fitness.

$$p(S_i^k) = \frac{R(S_i^k)}{\sum\limits_{i=1}^{n} R(S_i^k)}. \tag{28}$$

The cumulative probability $accp(S_i^k)$ is obtained by adding the probabilities.

$$accp(S_i^k) = \sum_{j=1}^{i} p(S_i^k). \tag{29}$$

Then, find the cumulative probability greater than the random number, and inherit the individual at the position where the first cumulative probability greater than the random number is.

4) CODING

Here we use binary encoding to encode the selected individuals, so that each individual is expressed as a chromosome in the genetic space. In this way the genetic algorithm can deal with the considered problem. We conduct the following processing for each selected individual:

$$q(S_i^k) = S_i^k \times 10^6. \tag{30}$$

Then convert $q(S_i^k)$ to binary code, which is represented by $S_i^{'k}$. Here we set the encoding length to L.

5) CROSSOVER

In this step, the selected individuals are crossed to generate new individuals of the next generation. First, we determine the parents of the cross, which are represented by $S_i^{'k}$ and S_i^{*k} respectively. Then the same cross position r is randomly determined on the binary code of the parents. We randomly generate a random number. If the random number is less than the hybridization rate, let the parents hybridize at the cross position r. Exchanged the code of the parents after the cross position r. The formula is as follows:

$$S_i^{'k}(L-r) = S_i^{*k}(L-r). \tag{31}$$

After crossing, the new individuals generated are represented by $S_i^{''k}$ and S_i^{**k}. Here we will inherit $S_i^{''k}$ as the next generation. The new individuals of the next generation after cross generation are represented by $Q^{'k+1}$, that is

$$Q^{'k+1} = S_i^{''k}, \tag{32}$$

where $Q^{'k}$ represents the kid produced by hybridization.

6) MUTATION

In the process of inheritance, mutation occurs by chance. In $Q^{'k+1}$ obtained in the previous step, the mutation position s is randomly generated. Randomly generate a random number. If the random number is less than the mutation rate, the mutation occurs. We express the mutation process by changing the coding value of the mutation position as follows

$$Q^{'k+1}(s) = 1 \overset{mutation}{\longleftrightarrow} 0. \tag{33}$$

7) DECODING

At this point, what we get is the code of each individual in the new population after a generation of inheritance. We decode the binary code to restore the parameter values of the population.

IV. SIMULATION RESULTS AND DISCUSSION

In the simulation, we assume that the maximum frequency hopping period $T_{\max}$ is 50 ms, and the frequency switching time T_0 is 1 ms. The maximum transmit power of the transmitter is $P_{smax} = 2$ W, and the maximum transmit power of the jammer is $P_{jmax} = 5$ W. The noise power $\delta^2 = -50$ dBm. The number of optional channels M is 32. In the genetic algorithm, the initial population size is 100, the coding length is 24, the selection rate is 0.5, the hybridization rate is 0.7, and the mutation rate is 0.001. The simulation adopts the outdoor scenario, and the simulation scenario is shown in Fig. 3. The distance between the jammer J and the receiver D is 6 km. The transmitter S moves on the dashed line along the direction of the arrow. In the simulation process, we first consider the scenario where the transmitter is 4 km away from the jammer and the receiver. Then we compare the anti-jamming performance of different schemes when the transmitter S moves.

Inspired by the path-loss model in [22] which has been widely used in wireless communications, the channel gains of S-D, S-J and J-D are respectively denoted as $h_{sd} = K(d_0/d_{sd})^{\gamma}$, $h_{sj} = K(d_0/d_{sj})^{\gamma}$ and $h_{jd} = K(d_0/d_{jd})^{\gamma}$, where K is a coefficient that depends on antenna characteristics and average channel loss, d_0 is the reference distance of antenna far field, γ is the path-loss factor, d_{sr}, d_{sj} and d_{jd} denote the distance of S-D, S-J and J-D respectively. The channel parameters are set as $K = 1$, $d_0 = 0.1$ km and $\gamma = 3$.

Assume that the transmitter observes two possible channel states of the channel gains between the jammer and the receiver. Because the channel gain is represented by the path loss model, the uncertainty of channel state information is caused by the uncertainty of the jammer position. Two possible jammer locations are: Case 1: $d_{sd} = 4$ km, $d_{sj} = 4$ km and $d_{jd} = 6$ km; Case 2: $d_{sd} = 4$ km, $d_{sj} = 3.5$ km and $d_{jd} = 5.5$ km. We got the simulation results for the perfect

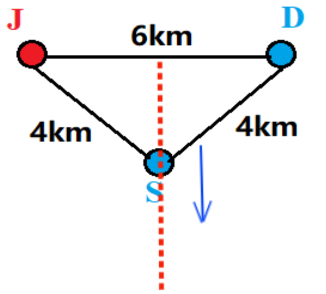


FIGURE 3. The simulation scenario.

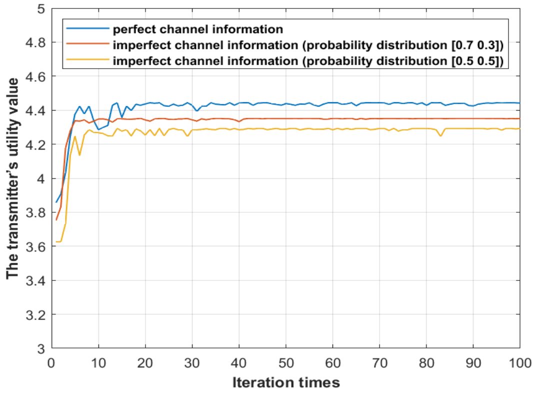


FIGURE 4. The convergence of the transmitter's utility value.

channel information and imperfect channel information. For the perfect channel information we can regard the probability distribution of the two possible channel gains is [1, 0]. For the imperfect channel information, we consider two different channel uncertainties where the probability distributions of the two channel states in Case 1 and Case 2 are [0.7, 0.3] and [0.5, 0.5] in the simulations.

The convergence performance of the transmitter's utility is given in Fig. 4. It can be seen that with perfect channel information, the transmitter's utility value is basically stable at around 4.44. When channel uncertainties increase, the converged value of the transmitter's utility decreases.

Fig. 5 shows the searched optimal power at S when different degrees of channel uncertainties are considered. It can be seen that the optimal transmit power of the transmitter is 0.35 W when perfect channel information is available, that is, $P_s{}^* = 0.35$ W. Bringing $P_s{}^*$ into the closed-form solutions of T^* and $T_E{}^*$, we can get $T^* = 0.0046$ s, $T_E{}^* = 0.0022$ s. When imperfect channel information is available, we can see from Fig. 5 the optimal transmit power of the transmitter is close to 0.35 W, that is, the uncertainty of channel state information does not influence the optimal power much.

The effect of the maximum jamming power on the transmitter's utility value is shown in Fig. 6. It can be seen that the utility value of the transmitter is not sensitive to the maximum jamming power, especially when the maximum jamming power becomes larger. The reason is that if the jamming power is much larger than the transmit power, the SINR at D becomes quite small, and larger jamming power will bring insignificant changes to the SINR at D, which means lager jamming power will bring limited influence on the transmitter's utility. If the received SINR is quite small, the transmitter will choose frequency hopping to deal with the jamming, instead of blindly increasing its transmit power. It is obvious in Fig. 6 that even the maximum jamming power keeps increasing, the transmitter always choose the transmit power around 0.35 W.

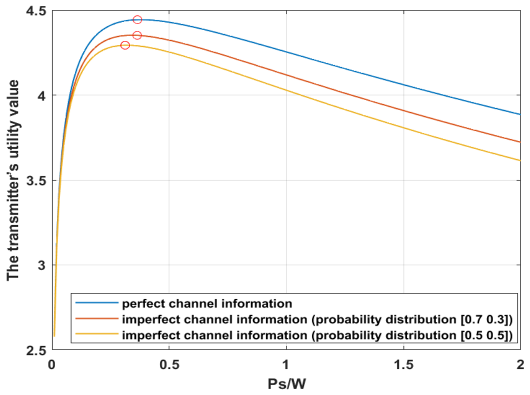


FIGURE 5. The optimal transmit power under different channel uncertainties.

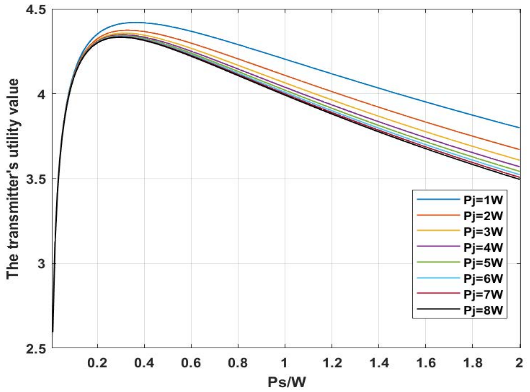


FIGURE 6. Influence of the maximum jamming power on the transmitter's utility value.

In Fig. 7, the influence of the maximum jamming power on the transmitter's frequency hopping period and the jammer's signal detection time is discussed. In this simulation, the transmit power is set to be 0.35 W. First, we can see the transmitter's frequency hopping period and the jammer's signal detection time are always positive. As can be seen from Fig. 7, as the maximum jamming power increases, the optimal frequency hopping period decreases, which means the transmitter will hop to another frequency band more frequently when the jammer increases the jamming power.

In Fig. 8, the multi-domain random scheme, the power-domain game scheme and the time-domain game scheme are compared with the proposed multi-domain game scheme. Specifically, the comparisons are performed under the perfect and imperfect channel gains in Fig. 8-a and Fig. 8-b respectively. In the time-domain game scheme, the transmit power P_s is randomly selected, the optimal frequency hopping period T is obtained by traversal search, and the optimal signal detection period T_E is selected

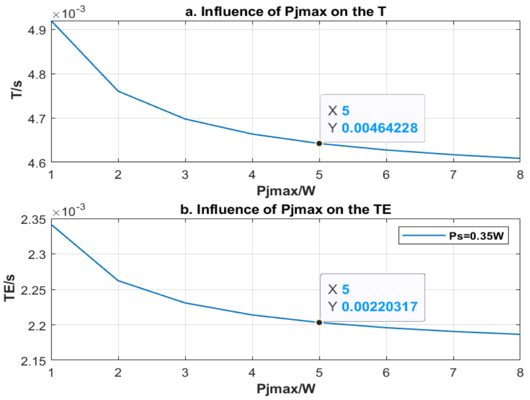


FIGURE 7. Influence of the maximum jamming power on the transmitter's frequency hopping period and the jammer's signal detection time.

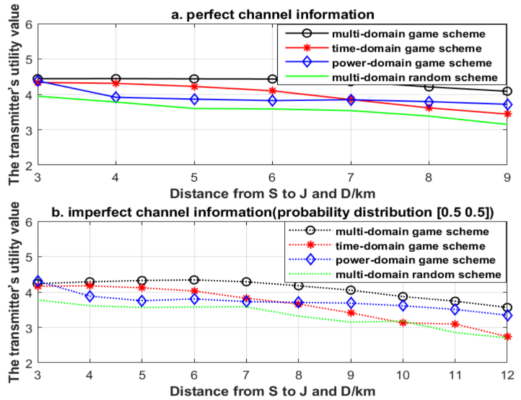


FIGURE 8. Performance comparison of the transmitter's utility for different schemes.

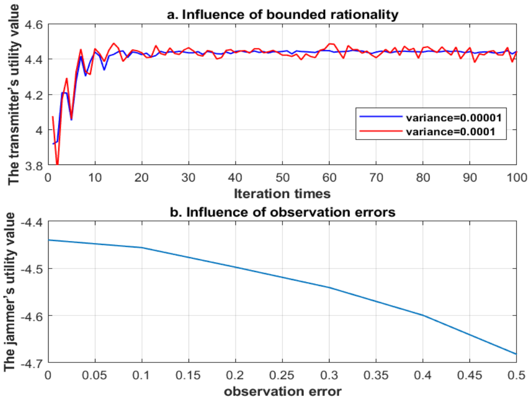


FIGURE 9. Influence of bounded rationality and observation errors.

by the closed-form solution obtained in this paper. In the power-domain game scheme, the frequency hopping period is randomly selected, the optimal transmit power P_s is obtained by traversal search, and the optimal sensing time T_E is derived by using the closed-form solution in this paper. The multi-domain random scheme is to randomly select the time domain and the power domain parameters. In the simulation, we also consider different locations of the transmitter. We let the transmitter move on the vertical line of the connection between the jammer and the receiver, starting from the midpoint of the connection, to a position of 12 km away from the jammer and the receiver. The performances of the schemes are compared for different locations of S. By observing Fig. 8, we can find that the proposed multi-domain scheme always obtains the largest utility value compared to the single-domain game schemes and the multi-domain random scheme no matter the channel information is perfect or not.

The influence of the imperfect information including the jammer's bounded rationality and observation errors is investigated in Fig. 9. In Fig. 9-a, we can see that the bounded rationality causes fluctuations to the transmitter's utility curve, the more irrational the jammer is, the more severe the curve jitters. But the bounded rationality does not change the trend of the curve, and the average converged value is always around 4.4 no matter the degree of the jammer's rationality. This observation is consistent with the result obtained in [25], which showed that bounded rationality will not have a great impact on the game model in general. Therefore, the effect of the jammer's bounded rationality on the transmitter utility is negligible when the iteration converges.

As for the curves in Fig. 9-b, the parameter setting in Fig. 9-b is not related to the parameters in Fig. 8-a. It can be seen that the larger the observation errors are, the smaller utility value the jammer can obtain. The reason is that, when the jammer is solving its optimal strategy, the jammer's observation errors of the transmitter's strategy bring difficulties to the jammer to find the optimal strategy, which results in the reduction of the jammer's utility.

V. CONCLUSION

In this paper, we propose a multi-domain anti-jamming game scheme to deal with a smart jammer which can sensing the legitimate communication and adjust its jamming policy. When channel uncertainties are considered, Bayesian Stackelberg game is used to model the competition between the transmitter and the jammer, where the transmitter plays the role of leader and the jammer acts as a follower. The imperfect information is modeled when formulating the utility functions. By using backward induction, the optimal time-power domain strategies for the transmitter and the jammer are obtained. The simulation results show that the proposed multi-domain game scheme obtains the largest utility value for the transmitter compared to the single-domain game schemes and multi-domain random scheme in different scenarios. The effect of the maximum jamming power is analyzed, and the influences of the imperfect information including observation errors, the bounded rationality and the channel uncertainties are discussed. We can see that the channel uncertainties can degrade the transmitter's utility, but the jammer's bounded rationality has insignificant impact on the transmitter's utility.

REFERENCES

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

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

[3] L. Jia, F. Yao, Y. Sun, Y. Niu, and Y. Zhu, "Bayesian Stackelberg game for antijamming transmission with incomplete information," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 1991–1994, Oct. 2016.

[4] Y. Li, L. Xiao, J. Liu, and Y. Tang, "Power control Stackelberg game in cooperative anti-jamming communications," in *Proc. 5th Int. Conf. Game Theory Netw.*, Beijing, China, 2014, pp. 1–6.

[5] 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.

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

[7] E. Ho, A. Rajagopalan, A. Skvortsov, S. Arulampalam, and M. Piraveenan, "Game theory in defence applications: A review," *Sensors*, vol. 22, no. 3, p. 1032, 2022.

[8] 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 Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2337–2352, May 2015.

[9] D. Yang, G. Xue, J. Zhang, A. Richa, and X. Fang, "Coping with a smart jammer in wireless networks: A Stackelberg game approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4038–4047, Aug. 2013.

[10] S. D'Oro, E. Ekici, and S. Palazzo, "Optimal power allocation and scheduling under jamming attacks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 3, pp. 1310–1323, Jun. 2017.

[11] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[12] Y. Xu, G. Ren, J. Chen, L. Jia, and Y. Xu, "Anti-jamming transmission in UAV communication networks: A Stackelberg game approach," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Qingdao, China, Oct. 2017, pp. 1–6.

[13] H. Noori and S. S. Vilni, "Defense against intelligent jammer in cognitive wireless networks," in *Proc. 27th Iranian Conf. Electr. Eng. (ICEE)*, Yazd, Iran, 2019, pp. 1309–1314.

[14] Y. Xu, J. Wang, Q. Wu, J. Zheng, L. Shen, and A. Anpalagan, "Dynamic spectrum access in time-varying environment: Distributed learning beyond expectation optimization," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5305–5318, Dec. 2017.

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

[16] G. Han, L. Xiao, and H. V. Poor, "Two-dimensional anti-jamming communication based on deep reinforcement learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, New Orleans, LA, USA, Mar. 2017, pp. 2087–2091.

[17] 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.

[18] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *Proc. 12th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, Hammamet, Tunisia, 2014, pp. 247–254.

[19] 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.

[20] Y. Zhang and F. Q. Yao, "The research on impairment of frequency hopping communication systems," *J. Xidian Univ.*, vol. 32, no. 6, pp. 472–476, 2005.

[21] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.

[22] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[23] J. Pita, M. Jain, M. Tambe, F. Ordóñez, and S. Kraus, "Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition," *Artif. Intell.*, vol. 174, no. 15, pp. 1142–1171, 2010.

[24] 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 Commun. Lett.*, vol. 6, no. 6, pp. 818–821, Dec. 2017.

[25] J. Yu and W. Jia, "Game model in the study of bounded rationality," *Scientia Sinica Math.*, vol. 50, no. 9, p. 1375, 2020.

[26] K. Yao, Y. Luo, Y. Yang, X. Liu, Y. Zhang, and C. Yao, "Location-aware incentive mechanism for traffic offloading in heterogeneous networks: A Stackelberg game approach," *Entropy*, vol. 20, no. 4, p. 302, 2018.

[27] S. Chen, H. Chen, and S. Jiang, "Optimal decision-making to charge electric vehicles in heterogeneous networks: Stackelberg game approach," *Energies*, vol. 12, no. 2, p. 325, 2019.

YONGCHENG LI received the master's degree from the University of Science and Technology of China, in 2012.

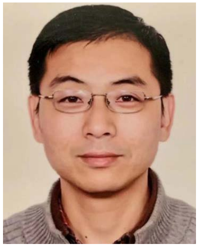
Then, he joined the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. His research interests include cognitive radio networks and comprehensive effect mechanism of electromagnetic environment.

KANGZE LI received the B.E. degree in communication engineering from Xi'an Jiaotong University, Xi'an, China, in 2020, where he is currently pursuing the master's degree in information and communication engineering.

His current research interest includes anti-jamming techniques in wireless communication systems.

ZHENZHEN GAO received the B.S. degree in communication engineering from Lanzhou University, Lanzhou, China, in 2005, and the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2011.

From August 2009 to September 2011, she was a Visiting Student with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. Since 2012, she has been with the School of Information and Communication Engineering, Xi'an Jiaotong University, where she is currently an Associate Professor. She is also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. Her current research interests include physical-layer security, index modulation, and advanced techniques in 5/6G wireless communication networks.



CHUNLEI ZHENG is currently with the Shanghai Institute of Microsystem and Information Technology Chinese Academy of Sciences. His current research interests include sensor signal processing, micro energy collection, low-power wireless passive sensor networks, and the Internet of Things.