Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 2786314, 14 pages
https://doi.org/10.1155/2020/2786314

Research Article

Adaptive Secure MIMO Transmission Mechanism against Smart Attacker

Xujun Shen, Qingchun Chen, Yulong Nie, and Keming Gan

Department of Electronics and Communication Engineering, Guangzhou University, Guangzhou 510006, China

Correspondence should be addressed to Qingchun Chen; qcchen@gzhu.edu.cn

Received 30 July 2019; Accepted 8 January 2020; Published 10 February 2020

Academic Editor: Daniele Pinchera

The MIMO transmission against a smart attacker has recently been formulated as a noncollaborative game, in which both the MIMO transmitter and the malicious attacker try to maximize their predefined utilities. In this paper, by carefully analyzing the Nash Equilibrium (NE), we focus on the conditions, in which the gaming results incline to the malicious attacker instead of the MIMO transmitter. In this adverse case, it is highly desirable to develop an effective mechanism to suppress the attack intention by the attacker for better secure communication. Motivated by this, an adaptive secure MIMO transmission scheme was proposed to make the MIMO transmitter better resist malicious attackers in adverse channel conditions. Compared with the existing gaming-based strategy, not only the transmit power of the MIMO transmitter but also the transmission probability will be adjusted in the proposed adaptive secure transmission scheme. Our analysis results show that the proposed scheme can be regarded as a generalized adaptive transmission one, i.e., when the adaptive transmit power policy is enough to suppress the attack motivation, the proposed scheme will be reduced to the adaptive power control scheme; otherwise, both the adaptive transmit power and the adaptive probabilistic transmission can be employed to suppress the attack motivation. The analysis results confirm us that the proposed adaptive transmission scheme provides us a choice to enhance the secure MIMO transmission performance in adverse conditions.

1. Introduction

When a malicious attacker can cleverly switch its attack mode amongst eavesdropping [1], jamming [2], and spoofing [3] to obstruct the secure communication between a transmitter and its receiver, it will impose critical challenges on the transmission strategy design for secure transmission. Game theory provides us a useful framework to derive the optimal transmission strategy in the presence of uncertain attack modes [4–6]. A MIMO wiretap zero-sum game was formulated in [4] to assume the secrecy rate as the utility function to analyze the conditions of equilibrium outcomes with various strategies. An exemplary multichannel spectrum access game (SAG) with unknown environment dynamics and limited information of other players is considered in [5] to find the best communication strategy through the joint reinforcement learning and type identification algorithm. In [6], a Bayesian game theory model is presented to allocate the defense effort among the nodes, in which the network defenders have little knowledge of the opponent's information and an explicit solution to this dilemma is derived. Afterwards, game theory-based secure transmission was extensively studied in [7–10] to cope with active attackers (smart attackers). A zero-sum game between the MIMO transmitter and the jamming-capable eavesdropper was formulated in [7] to derive the optimal mixed strategy and the realized Nash equilibria. By using game-theoretical methods, the secret communication in the multichannel network was studied in [8]. In [9], two simple stochastic games are proposed to deal with the dual threat of jamming and eavesdropping. The results show that, under some conditions, incorporating time slot for malicious threat detection in the transmission protocol can improve the confidentiality and reliability of the communication without additional transmission delay.

Motivated by the aforementioned progresses, some research efforts strived to combine the game theory with reinforcement learning to devise the secure transmission strategy. In [10], the Q-learning-based strategy was developed to cope with the smart attacker. The reinforcement learning-based spoofing detection and the deep reinforcement learning-based authentication scheme were proposed to further enhance the authentication in [11]. The Q-learning was utilized. The reinforcement learning-based power control scheme was developed in [12] to resist jamming attacks for the communication between the in-body sensors and the WBAN coordinator. In [13], the interaction between the user and a smart interferer in an ambient backscatter communication network was formulated as a game, and the closed-form equilibrium of the Stackelberg game was obtained. Due to lack of information about the system SNR and transmission strategy of the interferer, the Q-learning algorithm was proposed to derive the optimal strategy in a dynamic iterative manner. A zero-sum game between a base station equipped with multiple antennas and a smart jammer in the nonorthogonal multiple access (NOMA) system was formulated in [14]. The Stackelberg equilibrium of the antijamming NOMA transmission game was derived, and the reinforcement learning-based power control scheme was proposed for the downlink NOMA transmission without being aware of the jamming and radio channel parameters.

As a reinforcement learning (RL) algorithm to maximize the long-term expected reward in multistate environments, Q-learning provides us an effective technical solution for application in multiplayer general-sum games. The conditions that the Q-learning will converge with probability 1 to the optimal ones, the dynamic analysis framework, and the asymptotic convergence classes were addressed in [15–17], respectively. The MIMO transmission against a more powerful smart attacker, which can apply programmable radio devices (for instance, software-defined radios) to perform multiple types of attacks like eavesdropping, jamming and spoofing, was investigated in [18]. It is shown that the problem can be formulated as a noncooperative game, in which the power control strategy via reinforcement learning can be utilized to suppress the attack motivation of smart attackers in a dynamic MIMO transmission game without being aware of the attack and the radio channel model. Nonetheless, it should be addressed that most of the existing research efforts focus on how to derive the optimal power control strategy for suppressing the attack motivation of the malicious attacker by employing the reinforcement learning. Unfortunately, the power control strategy only will become no longer effective in terms of suppressing the attack motivation by the malicious attacker, especially in an adverse environment, as we will illustrate in this paper. In this case, it will become highly desirable to develop a new mechanism to make the MIMO transmitter better resist malicious attacker. And this is exactly the most important research motivation of our work in this paper.

Nash Equilibrium (NE) provides us the basis in the noncooperative game framework to determine the optimal solution, in which each player lacks any incentive to change his/her initial strategy, because a player does not gain anything by deviating from the initially chosen strategy, while other players keep their strategies unchanged. The NE analysis in [18] was presented to unveil that, when the transmit power of the MIMO transmitter is selected to be large enough, the game between the MIMO transmitter and the smart malicious attacker will incline to the MIMO transmitter, namely, the attacker tends to be idle. Although the proposed power control strategy via reinforcement learning can be utilized to cope with the smart attacker, one may readily observe the limitation of this design. In some adverse channel conditions, a very large transmit power might be infeasible, especially when the transmit power is limited. In this case, the challenges of the secure MIMO transmission arise. Exploring the critical factors that affect the strategy of both parties and developing an effective scheme to make the MIMO transmitter realize secure transmission against the attacker under an adverse channel condition are the other two motivations of our work in this paper. To this end, an adaptive secure MIMO transmission scheme was proposed in this paper, in which not only the transmit power of the MIMO transmitter but also the transmission probability will be adjusted in the proposed adaptive secure transmission scheme. And the proposed scheme can be interpreted as a generalized adaptive transmission strategy. When the adaptive transmit power policy is enough to suppress the attack motivation, the proposed scheme will be reduced to the adaptive power control scheme; otherwise, both the adaptive transmit power and the adaptive probabilistic transmission will be employed to suppress the attack motivation of the smart attacker. The contributions of this paper can be briefly summarized as follows:

(i) A comprehensive *Nash Equilibrium* (NE) analysis of the noncollaborative game framework in [18] was presented to explore the critical factors that dominate the game decisions. In this way, we show that the power control strategy only is not enough if we wish to suppress the attack motivation by the malicious attacker in adverse channel conditions.

(ii) A new probabilistic transmission scheme was proposed to realize a novel adaptive secure MIMO transmission scheme against the smart attacker. It is shown that, with both the probabilistic transmission control and the power control policy, we can improve the capability of the MIMO transmitter to resist the smart attacker, especially when compared with the original scheme in [18].

The remainder of this paper is organized as follows: The system model and the game problem formulation between the MIMO transmitter and the malicious attacker will be reviewed in Section 2. The comprehensive NE analysis will be presented in Section 3 to show that the game decision will not always be friendly to the MIMO transmitter. The proposed adaptive secure MIMO transmission mechanism will be addressed in Section 4. Numerical analysis was presented in Section 5 to show the applicability of the proposed scheme. Finally, we conclude our work in Section 6.

Notations. All boldface letters indicate vectors (lower case) or matrices (upper case). The superscripts $(\cdot)^H$ denote the conjugate transpose. $\det(\cdot)$ represents the determinant of a matrix. $\mathbf{I}_k$ is a $k \times k$ identity matrix.

2. System Model and the MIMO Game Problem Formulation

2.1. System Model. Let us consider the *Alice-Bob-Eve* communication model illustrated in Figure 1, in which *Alice* is assumed to be provisioned with M transmit antennas, *Bob* is assumed to be provisioned with N_r receive antennas, and *Eve* has N antennas. *Alice* is supposed to communicate with *Bob*, while *Eve* is smart and he will try to choose his attack mode amongst eavesdropping, jamming, and spoofing to obstruct the communication between *Alice* and *Bob*. The *Alice-Bob* link, the *Alice-Eve* link, and the *Eve-Bob* link can be represented by the $N_r \times M$ channel matrix $\mathbf{H}_{ba}$, the $N \times M$ channel matrix $\mathbf{H}_{ea}$, and the $N_r \times N$ channel matrix $\mathbf{H}_{be}$, respectively. Throughout the paper, the i-th largest eigenvalue of $\mathbf{H}_m\mathbf{H}_m^H$ will be denoted by λ_i^m, $m \in \{ba, be, ea\}$.

Alice is assumed to transmit an M-dimensional signal vector $\mathbf{x}_a$ to *Bob*, and the transmit power is $E[\mathbf{x}_a^H\mathbf{x}_a] = P, P \in [0, P_{\max}]$, where $P_{\max}$ stands for the maximally allowed transmit power constraint at *Alice*. When *Eve* decides not to attack the communication (namely, the attack mode indicator $q = 0$), the MIMO transmission rate R can be given by [19]

$$\begin{aligned} R &= \log_2 \det\left(I + \frac{P}{M}\mathbf{H}_{ba}\mathbf{H}_{ba}^H\right) \\ &= \sum_{i=1}^{\min\{M,N_r\}} \log_2\left(1 + \frac{P\lambda_i^{ba}}{M}\right). \end{aligned} \tag{1}$$

If *Eve* chooses to overhear the transmission by *Alice* (namely, the attack mode indicator $q = 1$), the received signal at *Eve* will be

$$\mathbf{y}_e = \mathbf{H}_{ea}\mathbf{x}_a + \mathbf{n}_e, \tag{2}$$

where $\mathbf{n}_e$ represents the N-dimensional additive zero-mean white Gaussian noise vector, i.e., $\mathbf{n}_e \sim \mathcal{CN}(0, \mathbf{I}_N)$. In this case, the achievable MIMO secrecy rate R_E will be [2]

$$\begin{aligned} R_E &= \log_2\det\left(I + \frac{P}{M}\mathbf{H}_{ba}\mathbf{H}_{ba}^H\right) - \log_2\det\left(I + \frac{P}{M}\mathbf{H}_{ea}\mathbf{H}_{ea}^H\right) \\ &= \sum_{i=1}^{\min\{M,N_r\}} \log_2\left(1 + \frac{P\lambda_i^{ba}}{M}\right) - \sum_{i=1}^{\min\{N,M\}} \log_2\left(1 + \frac{P\lambda_i^{ea}}{M}\right). \end{aligned} \tag{3}$$

If *Eve* decides to block Alice's transmission with the jamming signal $\mathbf{x}_J$ (namely, the attack mode indicator $q = 2$), and we assume $E[\mathbf{x}_J^H\mathbf{x}_J] = P_J$, P_J is the transmit power by *Eve* in the jamming mode, *Bob* will receive the following signal:

$$\mathbf{y}_J = \mathbf{H}_{ba}\mathbf{x}_a + \mathbf{H}_{be}\mathbf{x}_J + \mathbf{n}_b, \tag{4}$$

where $\mathbf{n}_b$ represents the N_r-dimension white Gaussian noise at *Bob*, i.e., $\mathbf{n}_b \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$. Then, the achievable secrecy rate R_J can be given by [2]

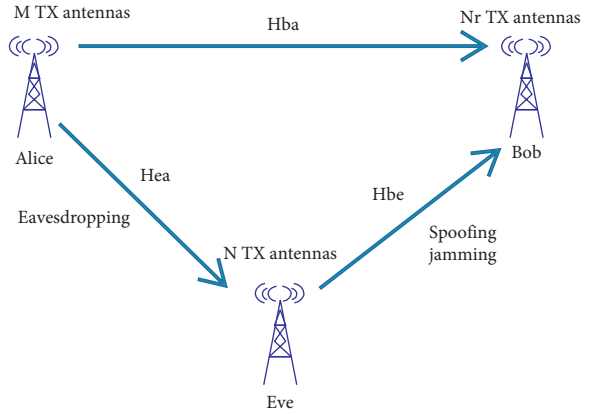


Figure 1: MIMO transmission in the presence of smart attacker.

$$\begin{aligned} R_J &= \log_2 \det\left(\mathbf{I} + \frac{P}{M}\mathbf{H}_{ba}\mathbf{H}_{ba}^H\left(\mathbf{I} + \frac{P_J}{N}\mathbf{H}_{be}\mathbf{H}_{be}^H\right)^{-1}\right) \\ &= \sum_{i=1}^{\min\{N,Nr\}} \log_2\left(1 + \frac{P\lambda_i^{ba}N}{M\left(N + \lambda_i^{be}P_J\right)}\right). \end{aligned} \tag{5}$$

If *Eve* chooses to spoof *Bob* by sending a fake signal $\mathbf{x}_s$ with restricted transmit power $E[\mathbf{x}_s^H\mathbf{x}_s] = P_s/N$ when *Alice* is silent (namely, the attack mode indicator $q = 3$), P_s is the transmit power by *Eve* in the spoofing mode and *Bob* will receive the following signal:

$$\mathbf{y}_s = \mathbf{H}_{be}\mathbf{x}_s + \mathbf{n}_b. \tag{6}$$

The attack mode of spoofing aims at transmitting spoofing information to *Bob* only. The achievable secrecy rate of *Alice* when *Eve* spoofs *Bob* can be given by [18]

$$\begin{aligned} R_S &= \log_2 \det\left(\mathbf{I} + \frac{P}{M}\mathbf{H}_{ba}\mathbf{H}_{ba}^H\right) - \gamma \log_2 \det\left(\mathbf{I} + \frac{P_S}{N}\mathbf{H}_{be}\mathbf{H}_{be}^H\right) \\ &= \sum_{i=1}^{\min\{M,Nr\}} \log_2\left(1 + \frac{P\lambda_i^{ba}}{M}\right) - \gamma \sum_{i=1}^{\min\{N,Nr\}} \log_2\left(1 + \frac{P_S\lambda_i^{be}}{N}\right), \end{aligned} \tag{7}$$

where γ represents the spoofing message utility coefficient.

2.2. MIMO Transmission Game Problem Formulation. In order to cope with the smart attacker, the MIMO transmission scheme can be derived by employing a noncollaborative game framework, in which the following two utility functions of *Alice* and *Eve* are assumed [18]:

$$\mu_a(P, q) = \ln 2 \cdot \sum_{k=0}^{K} U_k I(k = q) - C_a P, \tag{8}$$

$$\mu_e(P, q) = -\ln 2 \cdot \sum_{k=0}^{K} U_k I(k = q) - f(q), \tag{9}$$

where C_a represents the unit power consumption at *Alice*, $U_k \in \{R, R_E, R_J, R_S\}$, and $I(k = q) = 1$ if $k = q$; otherwise,

$I(k=q)=0$. $f(q)\in\{0,\theta_E,\theta_J,\theta_S\}$ represents the cost for *Eve* to choose attack mode $q\in\{0,1,2,3\}$. Considering the secure MIMO transmission game denoted by $G=\langle(Alice,Eve),(P,q),(\mu_a,\mu_e)\rangle$ with the game participants $(Alice,Eve)$, game strategy (P,q), and utility functions (μ_a,μ_e), the Nash Equilibrium (*NE*) strategy (P^*,q^*) should satisfy the following conditions:

$$\mu_a(P^*,q^*)\geq\mu_a(P,q^*),\quad \forall 0\leq P\leq P_{\max},\tag{10}$$

$$\mu_e(P^*,q^*)\geq\mu_e(P^*,q),\quad \forall q\in\{0,1,2,3\}.\tag{11}$$

The *NE* condition that is friendly to *Alice* and the issue on how to suppress completely the attack motivation by *Eve* are highlighted in [18], in which the reinforcement learning-based power control strategy was proposed as well to realize the attack-free game results (namely, $q=0$). Nonetheless, it should be addressed that there exist multiple *NE* conditions, in which the game results will not always incline to *Alice*. In this paper, our emphasis will be focused on the case, in which the game results incline to the malicious attacker instead of the MIMO transmitter. Obviously, it is highly desirable to develop an effective mechanism to cope with the smart attacker in this case for a better secure communication. In fact, this is exactly the problem that we would like to highlight in this paper.

3. Nash Equilibrium Analysis

3.1. Nash Equilibrium in support of Alice. Considering the utility functions in (8) and (9), as well as the *Nash Equilibrium* condition in (10) and (11), one may readily derive that, the *NE* condition at $(P,q)=(p^*,0)$ in support of *Alice* for the game $G=\langle(Alice,Eve),(P,q),(\mu_a,\mu_e)\rangle$ can be achieved when the following conditions are satisfied:

$$\begin{cases}\theta_E\geq\displaystyle\sum_{i=1}^{\min\{M,N\}}\ln\left(1+\frac{p^*\cdot\lambda_i^{ea}}{M}\right),\\[2ex]\theta_J\geq\displaystyle\sum_{i=1}^{\min\{N,Nr\}}\ln\left(1+\frac{p^*p_J\lambda_i^{ba}\lambda_i^{be}}{MN+Mp_J\lambda_i^{be}+p^*N\lambda_i^{ba}}\right),\\[2ex]\theta_S\geq\gamma\displaystyle\sum_{i=1}^{\min\{N,Nr\}}\ln\left(1+\frac{p_S\lambda_i^{be}}{N}\right).\end{cases}\tag{12}$$

In fact, three threshold values of θ_E, θ_J, and θ_S can be interpreted as the minimum reward expected by *Eve* if he decides to overhear *Alice*, to jam or to spoof *Bob*. The *NE* conditions in (12) implies that, when the realized attack reward by *Eve* is less than the predefined minimum expected reward, *Eve* will give up attacking, namely, the game between *Alice* and *Eve* inclines to *Alice*. One may readily observe that for the given expected minimum reward θ_E, θ_J, and θ_S, the achievability of the *NE* conditions at $(P,q)=(p^*,0)$ will depend on the underlying channel $\{\lambda_i^{ba},\lambda_i^{ea},\lambda_i^{be}\}$. One may readily note that, the above *NE* conditions cannot always be satisfied, especially when the channel $\{\lambda_i^{ba},\lambda_i^{ea},\lambda_i^{be}\}$ changes.

3.2. Nash Equilibrium in support of Eve. In the same way, on the basis of (10) and (11), we may derive the *Nash Equilibrium* conditions in support of *Eve* for $q=1,2,3$ in Table 1.

(1) Nash Equilibrium in Eavesdrop Mode. By comparing the *NE* conditions with those in (12), one may readily observe that the game decision will incline to the eavesdrop attack mode by *Eve* when the predefined eavesdropping reward θ_E can be fulfilled and the relative eavesdropping gain of $\Delta_E\triangleq\sum_{i=1}^{\min\{M,N\}}\ln(1+(p^*\lambda_i^{ea}/M))-\theta_E$ is the largest one among three attack modes. A further examination of the conditions unveils that better channel $\{\lambda_i^{ea}\}$ between *Alice* and *Eve* will make the game have a larger opportunity to incline to the eavesdrop attack mode, which complies with the heuristics, because now *Eve* is closer to *Alice* and he is in a better condition to overhear the transmission by *Alice*.

(2) The Nash Equilibrium in Jam Mode. In the same way, by comparing the *NE* conditions with that in (12), one may readily observe that the game decision will incline to the jam attack mode by *Eve* if the predefined jamming reward θ_E can be fulfilled and the relative jamming gain of $\Delta_J\triangleq\sum_{i=1}^{\min\{N,Nr\}}\ln(1+(p^*\cdot p_J\lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^*\cdot N\lambda_i^{ba})))-\theta_J$ is the best among three attack modes. By carefully examining the conditions, we may observe that reasonable channel $\{\lambda_i^{ba}\}$ between *Alice* and *Bob* and reasonable channel $\{\lambda_i^{be}\}$ between *Eve* and *Bob* will make the game have a larger opportunity to incline to the Jam attack mode (one may see this from $(p^*\cdot p_J/((MN/\lambda_i^{ba}\lambda_i^{be})+(Mp_J/\lambda_i^{ba})+p^*\cdot(N/\lambda_i^{be}))))$, which complies with the heuristics because now both *Alice* and *Eve* are in good conditions to transmit to *Bob*.

(3) The Nash Equilibrium in Spoof Mode. Similarly, by comparing the *NE* conditions with that in (12), one may readily conclude that the game decision will incline to the spoof attack mode by *Eve* if the predefined spoofing reward θ_S can be fulfilled and the relative spoofing gain of $\Delta_S\triangleq\gamma\cdot\sum_{i=1}^{\min\{N,Nr\}}\ln(1+(p_s\lambda_i^{be}/N))-\theta_S$ is the best among three attack modes. A further examination of the conditions unveils that better channel $\{\lambda_i^{be}\}$ between *Eve* and *Bob* will make the gaming have a larger opportunity to incline to the spoof attack mode, which complies with the heuristics because now the *Eve* is closer to *Bob* and he is now in a good condition to spoof the reception by *Bob*.

In order to explicate more clearly the above *NE* conditions, let us summarize briefly their dependency on the underlying channel gains of $\{\lambda_i^{ba},\lambda_i^{ea},\lambda_i^{be}\}$.

From the perspective of attack, a large λ_i^{be} (*Eve-Bob* link is good) is a beneficial situation for *Eve* to choose either jam or spoof mode. If λ_i^{ba} also becomes large (*Alice-Bob* link is good), *Eve* would incline to jam. The *NE* in support of *Eve* may happen when both the *Alice-Bob* link λ_i^{ba} and the *Eve-Bob* link λ_i^{be} are not in good conditions, while we have a

Table 1: *NE* conditions in support of *Eve*.

Attack mode	Conditions
Eavesdropping mode, $q=1$	$\theta_E \le \sum_{i=1}^{\min\{M,N\}} \ln(1+(p^*\lambda_i^{ea}/M))$ $\sum_{i=1}^{\min\{M,N\}} \ln(1+(p^*\lambda_i^{ea}/M)) - \theta_E \ge \sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p^* p_J \lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^*\cdot N\lambda_i^{ba}))) - \theta_J$ $\sum_{i=1}^{\min\{M,N\}} \ln(1+(p^*\lambda_i^{ea}/M)) - \theta_E \ge \gamma\cdot\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p_s\lambda_i^{be}/N)) - \theta_S$
Jamming mode, $q=2$	$\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p^*\cdot p_J\lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^*\cdot N\lambda_i^{ba}))) - \theta_J \ge \sum_{i=1}^{\min\{M,N\}} \ln(1+(p^*\lambda_i^{ea}/M)) - \theta_E$ $\theta_J \le \sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p^* p_J \lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^* N\lambda_i^{ba})))$ $\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p^*\cdot p_J\lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^*\cdot N\lambda_i^{ba}))) - \theta_J \ge \gamma\cdot\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p_s\lambda_i^{be}/N)) - \theta_S$
Spoofing mode, $q=3$	$\gamma\cdot\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p_s\lambda_i^{be}/N)) - \theta_S \ge \sum_{i=1}^{\min\{M,N\}} \ln(1+(p^*\lambda_i^{ea}/M)) - \theta_E$ $\gamma\cdot\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p_s\lambda_i^{be}/N)) - \theta_S \ge \sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p^*\cdot p_J\lambda_i^{ba}\lambda_i^{be}/(MN+Mp_J\lambda_i^{be}+p^* N\lambda_i^{ba}))) - \theta_J$ $\theta_S \le \gamma\cdot\sum_{i=1}^{\min\{N,Nr\}} \ln(1+(p_s\lambda_i^{be}/N))$

reasonable *Alice-Eve* link λ_i^{ea}. In this case, the game decision may incline to make *Eve* overhear the transmission by *Alice*. The above discussion clearly tells us that the game framework between *Alice* and *Eve* will not always incline to *Alice*. When the wireless environment becomes adverse, such that *Eve* inclines to attack, the power control strategy only would be no longer effective to suppress *Eve*'s attack motivation. As we will address in Section 4, now the adaptive probabilistic transmission design can be utilized to discourage the attack motivation by *Eve*.

4. New Adaptive Secure Transmission Design

4.1. Probabilistic Transmission Policy. In Section 3, we have shown that, the *NE* at $(p^*, 0)$ can be achieved when all the possible attack rewards by *Eve* are less than the predefined minimum expected rewards θ_E, θ_J, and θ_S. In this case, a reinforcement learning-based power control strategy can be employed to approach the *NE* at $(p^*, 0)$ [18]. The paid cost may be some increase in the required transmit power at *Alice*. However, when some attack reward by *Eve* is larger than the predefined minimum expected rewards, in this case, the transmit power control strategy may be no longer effective in terms of the attack suppression. In order to better fulfill the secure transmission requirements and to realize reasonable secure transmission in adverse conditions, we proposed to use the probabilistic transmission strategy. Unlike the original transmission scheme [18], in which *Alice* will always transmit irrespective of the current channel conditions, in the proposed adaptive probabilistic transmission scheme, *Alice* may decide to transmit in a probabilistic manner, especially when the game decision inclines to *Eve*. To this end, the utility function of *Alice* can be modified as follows:

$$\mu_a(\mathscr{P}_T, P, q) = \mathscr{P}_T \cdot \left(\ln 2 \cdot \sum_{k=0}^{K} U_k I(k=q) - C_a P \right), \tag{13}$$

where $\mathscr{P}_T \in (0,1]$ represents the transmission probability by *Alice*. One may readily observe that the above utility function will be reduced to the traditional utility function assumed in [18] when $\mathscr{P}_T = 1$. We will show later that the proposed adaptive probabilistic transmission scheme will subsume the original adaptive transmission scheme in [18] as a special case by letting $\mathscr{P}_T = 1$.

Because *Eve* will try to obstruct the communication between *Alice* and *Bob* with a certain paid cost, we may assume the worst case that *Eve* knows the probabilistic transmission control mechanism at *Alice*; the utility function at *Eve* can thus be revised as follows:

$$\mu_e(p_T, P, q) = -\mathscr{P}_T \cdot \ln 2 \cdot \sum_{k=0}^{K} U_k I(k=q) - f(q). \tag{14}$$

On the basis of the two updated utility functions in (13) and (14), we may derive the *NE*s in Table 2.

By comparing Tables 1 and 2, we may readily observe from the modified *NE* conditions how the proposed probabilistic transmission scheme is able to further suppress the attack motivations by *Eve*. By introducing the transmission probability control mechanism, now three possible attack rewards and the associated relative attack rewards will be reduced with a discount coefficient, which is proportional to the transmission probability $\mathscr{P}_T < 1$. And this explicates the philosophy that the proposed probabilistic transmission leads to the suppression of the attack motivation by *Eve*.

Of course, it should be addressed that the proposed probabilistic transmission scheme will incur degradation in the achieved secrecy capacity, in that now *Alice* will not always attempt to transmit, no matter how *Eve* reacts and what about the underlying channel conditions. Nonetheless, as we will illustrate in Section 5, in adverse channel conditions where the traditional game strategy (power control strategy) fails in suppressing the attack motivation by *Eve*, the proposed adaptive secure transmission scheme can still help to improve the secure transmission between *Alice* and *Bob* by discouraging the attack motivation by *Eve*. Thus, some loss in the transmission opportunity is still worthwhile.

4.2. Reinforcement Learning-Based Adaptive Secure Transmission Scheme. The *Q*-learning-based algorithm in [18] can be modified to derive both the optimal transmit power strategy and the optimal probabilistic transmission policy to realize the adaptive secure MIMO transmission scheme. As summarized in Algorithm 1, the *Q*-learning-based algorithm

Table 2: *NE* conditions with the probabilistic transmission scheme.

Attack mode	Conditions
No-attack mode, $q=0$	$\theta_E \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + p * (\lambda_i^{ea}/M))$ $\theta_J \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,N_r\}} \ln(1 + (p^* p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* N \lambda_i^{ba})))$ $\theta_s \geq \mathscr{P}_T \cdot \gamma \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_S \lambda_i^{be}/N))$
Eavesdropping mode, $q=1$	$\theta_E \leq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + (p^* \lambda_i^{ea}/M))$ $\mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + (p^* \lambda_i^{ea}/M)) - \theta_E \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p^* p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* \cdot N \lambda_i^{ba}))) - \theta_J$ $\mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + (p^* \lambda_i^{ea}/M)) - \theta_E \geq \mathscr{P}_T \cdot \gamma \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_s \lambda_i^{be}/N)) - \theta_S$
Jamming mode, $q=2$	$\mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p^* \cdot p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* \cdot N \lambda_i^{ba}))) - \theta_J \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + (p^* \lambda_i^{ea}/M)) - \theta_E$ $\theta_J \leq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p^* p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* N \lambda_i^{ba})))$ $\mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p^* \cdot p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* \cdot N \lambda_i^{ba}))) - \theta_J \geq \mathscr{P}_T \cdot \gamma \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_s \lambda_i^{be}/N)) - \theta_S$
Spoofing mode, $q=3$	$\mathscr{P}_T \cdot \gamma \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_s \lambda_i^{be}/N)) - \theta_S \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{M,N\}} \ln(1 + (p^* \lambda_i^{ea}/M)) - \theta_E$ $\mathscr{P}_T \cdot \gamma \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_s \lambda_i^{b})) - \theta_S \geq \mathscr{P}_T \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p^* \cdot p_J \lambda_i^{ba} \lambda_i^{be} / (MN + M p_J \lambda_i^{be} + p^* N \lambda_i^{ba}))) - \theta_J$ $\theta_S \leq \mathscr{P}_T \cdot \gamma \cdot \sum_{i=1}^{\min\{N,Nr\}} \ln(1 + (p_s \lambda_i^{be}/N))$

can be utilized to derive both the optimal power control p^* and the optimal probabilistic transmission $\mathscr{P}_T^*$ for *Alice*. Specifically, $Q(P, \mathscr{P}_T, s)$ stands for the Q function of *Alice*, in which s represents the system state and P and $\mathscr{P}_T$ denote two actions by *Alice*. $V(s)$ indicates the maximum of $Q(P, \mathscr{P}_T, s)$ over all possible actions, given the state of s. The learning rate $\alpha \in [0, 1]$ represents the weight of the current quality during the learning process, while $\delta \in [0, 1]$ is the discount factor that denotes the uncertainty of *Alice* about the future gains. *Alice* would observe the strategy by *Eve* in the $(n-1)$-th slot q^{n-1} and can assume it as its current state $s^* = q^{n-1}$. With the times going by, *Alice* is able to choose the optimal power control strategy $p^* \in (0, P_{\max}]$ and $\mathscr{P}_T^* \in (0, 1]$. In practical applications, we may consider to set a smallest probabilistic transmission parameter $\mathscr{P}_{T,\min}$ according to the least transmission rate requirement by *Alice-Bob* link.

5. Numerical Analysis

In order to show the applicability of the proposed adaptive probabilistic transmission scheme, we will focus on the conditions in which the traditional game results will incline to *Eve*. By doing so, we will show clearly that when the conditions tend to be in support of *Eve*, the traditional game will fail in guaranteeing secure transmission from *Alice* to *Bob*. On this basis, we then highlight how the proposed probabilistic transmission policy can be utilized along with the power control strategy to improve the secrecy transmission from *Alice* to *Eve* by suppressing the attack motivation of *Eve*. In all simulations, without statement, we assume $M = 5$, $N = 3$, $N_r = 3$, $\Theta = [2.2, 3, 3.2]$, $P_J = 3$, $P_s = 3.2$, and $P_{\max} = 10$. In simulations, we assume $C_a = 0.1$ and $\gamma = 0.5$. $\mathscr{P}_T \in [0.3, 1]$ is assumed to make sure that a very low transmission rate can be avoided.

5.1. No Attack Benchmark System. And the following channel conditions are assumed in simulations as the benchmark setting: $\{\lambda_1^{ba} = 2.77, \lambda_2^{ba} = 1.81, \lambda_3^{ba} = 1.07\}$, $\{\lambda_1^{ea} = 1.15, \lambda_2^{ea} = 0.75, \lambda_3^{ea} = 0.44\}$, and $\{\lambda_1^{be} = 4.01, \lambda_2^{be} = 2.29, \lambda_3^{be} = 0.93\}$. In this setup, although the *Alice-Bob* link is not the best among the three links, the transmit power control strategy at *Alice* can successfully suppress the attack motivation by *Eve*, as shown in Figure 2(b). As illustrated in Figure 2(c), with some paid cost in the increased transmit power at *Alice*, secure transmission from *Alice* to *Bob* can be realized (see Figure 2(a)). For illustration purpose, the realized secrecy capacity, different attack mode probabilities by *Eve*, and the required transmit power at *Alice* in the same benchmark system are illustrated in Figure 3. One may readily observe that by introducing the probabilistic transmission control strategy, we can also guarantee the suppression of the attack motivation by *Eve* for a secure transmission from *Alice* to *Bob*. Meanwhile, less transmit power is required due to the probabilistic transmission strategy (Figure 3(c)). The paid cost is some loss in the achieved secrecy rate from *Alice* to *Bob*. Then, in the following numerical analysis, we will show that the proposed probabilistic transmission strategy can be utilized with the transmit power control to formulate a more robust transmission scheme in the presence of malicious attack.

5.2. Ability to Suppress the Eavesdropping. Let us consider the following channel conditions as a typical eavesdropping setup: $\{\lambda_1^{ba} = 2.77, \lambda_2^{ba} = 1.81, \lambda_3^{ba} = 1.07\}$, $\{\lambda_1^{ea} = 5.77, \lambda_2^{ea} = 3.78, \lambda_3^{ea} = 2.24\}$, and $\{\lambda_1^{be} = 2.00, \lambda_2^{be} = 1.14, \lambda_3^{be} = 0.46\}$. Compared with the benchmark system in Section 5.1, now the *Alice-Eve* link is the best among the three involved links, the *Eve-Bob* link is the worst, and the *Alice-Bob* channel remains unchanged. According to our analysis in Section 3.2, now the traditional game decision will incline to the eavesdropping by *Eve*. As shown in Figure 4(a), now the power control strategy will only fail in suppressing *Eve*'s motivation to overhear the transmission by *Alice*. In fact, the eavesdropping probability by *Eve* is now about 80%, as illustrated in Figure 4(b). We do not include the secrecy rate performance since now there is in fact no secure rate at all.

Obviously, it is highly desirable for *Alice* to figure out an effective mechanism to resist the attack by *Eve* to fulfill the secure transmission requirement. In the same eavesdropping setting, if the proposed adaptive probabilistic

(1) **Initialize** $q^0 = 0$, $Q(P, \mathscr{P}_T, s) = 0$, $V(s) = 0$, $\forall s, P, \mathscr{P}_T$.
(2) **For** $n = 1, 2, 3, \ldots$ **do**
(3) Update the state $s^n = q^{n-1}$
(4) Choose P^n and $\mathscr{P}_T^n$ with the ε-greedy policy
(5) Send signal with power P^n and probabilistic parameter $\mathscr{P}_T^n$ over M antennas
(6) Observe the attack type q^n and μ_a
(7) Update the Q function and value function:
(8) $Q(P^n, \mathscr{P}_T^n, s^n) = (1 - \alpha)Q(P^n, \mathscr{P}_T^n, s^n) + \alpha(\mu_a(P^n, \mathscr{P}_T^n, s^n) + \delta(V(s^{n+1})))$
(9) $V(s^n) = \max_{P, \mathscr{P}_T} Q(P, \mathscr{P}_T, s^n)$, $(0 \le P \le P_{\max}, \mathscr{P}_{T,\min} < \mathscr{P}_T \le 1)$
(10) **end for**

Algorithm 1: Adaptive secure MIMO transmission scheme via Q-learning.

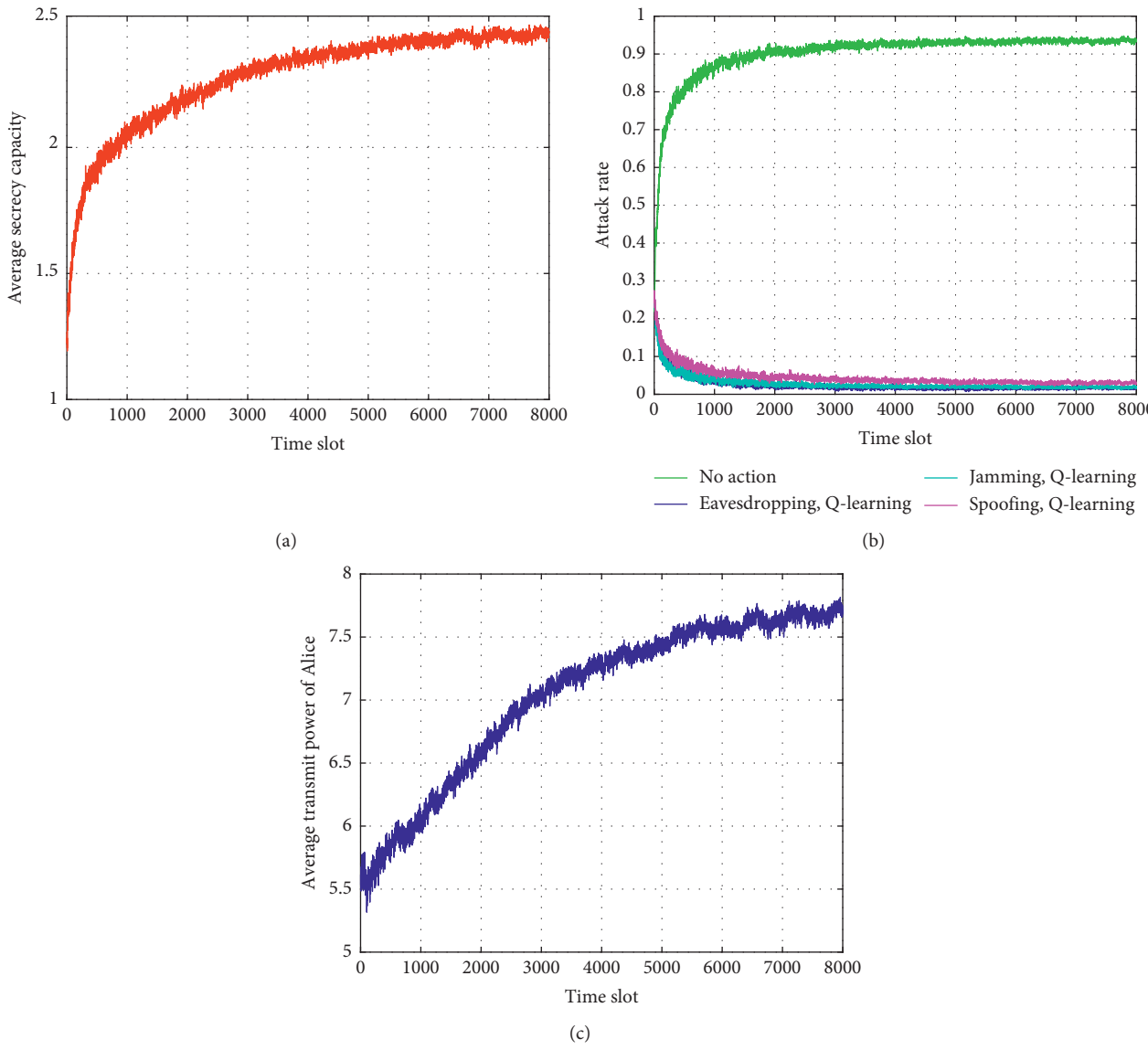


Figure 2: Illustration of the *NE* in the no-attack mode by employing the power control strategy in the benchmark system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power P.

transmission scheme is utilized along with the power control strategy, the realized secrecy rate, different attack mode probabilities by *Eve*, the required transmit power, and the probabilistic transmission control at *Alice* are illustrated in Figure 5. We can see the game results now incline to *Alice* again, as illustrated in Figure 5(b). One may also note that by lowering the transmission probability of *Alice*, we can effectively suppress the attack motivation by *Eve*. As a result,

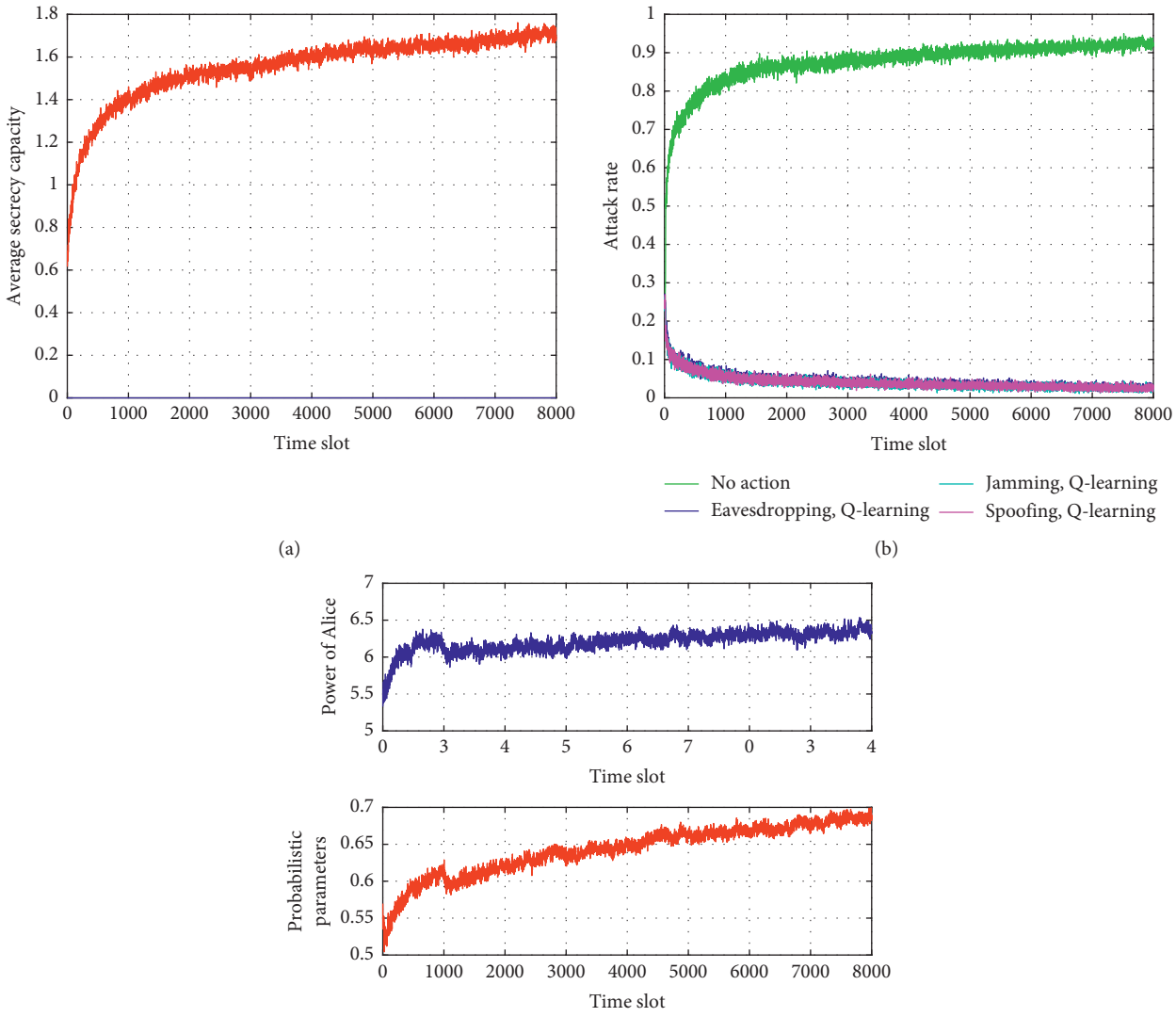


Figure 3: Illustration of the *NE* in the no-attack mode by employing both the power control strategy and the probabilistic transmission scheme in the benchmark system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit powers P and P_T.

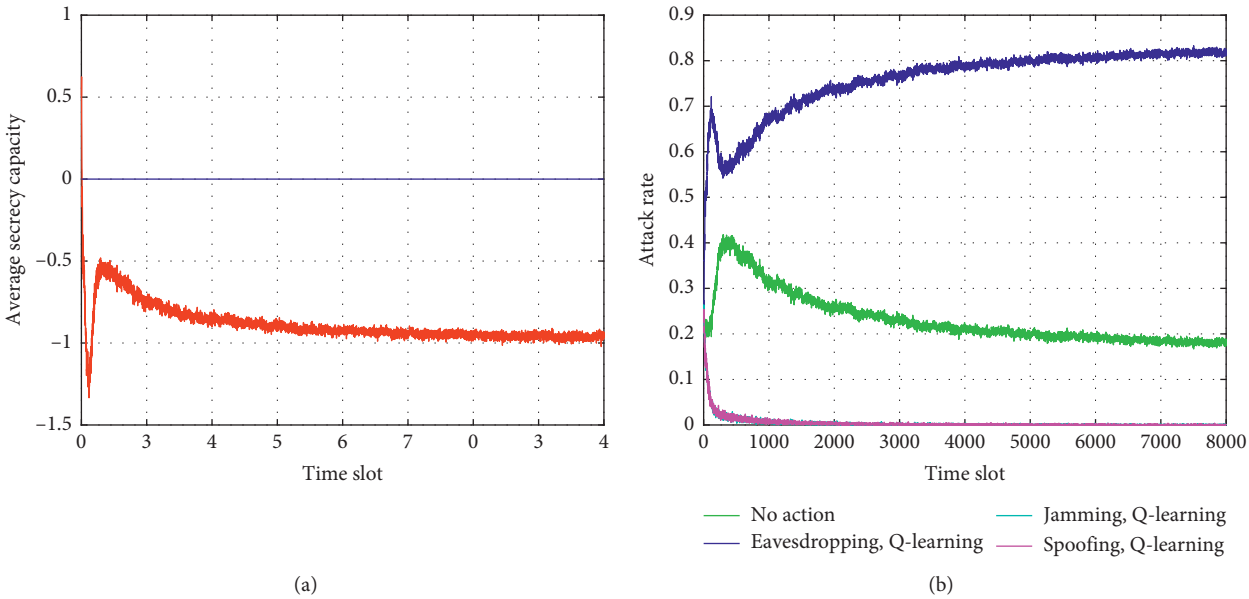
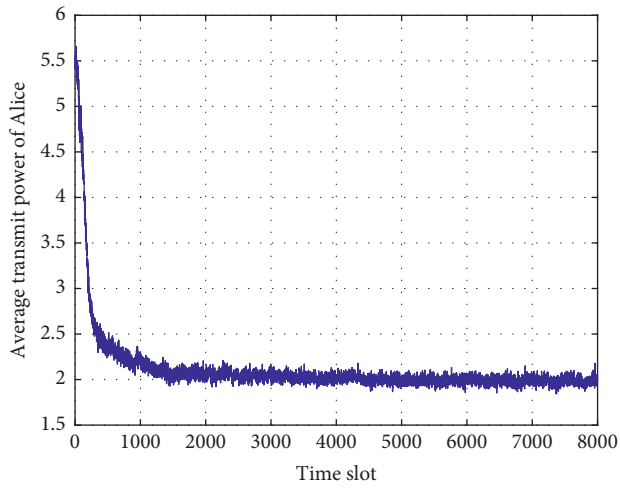
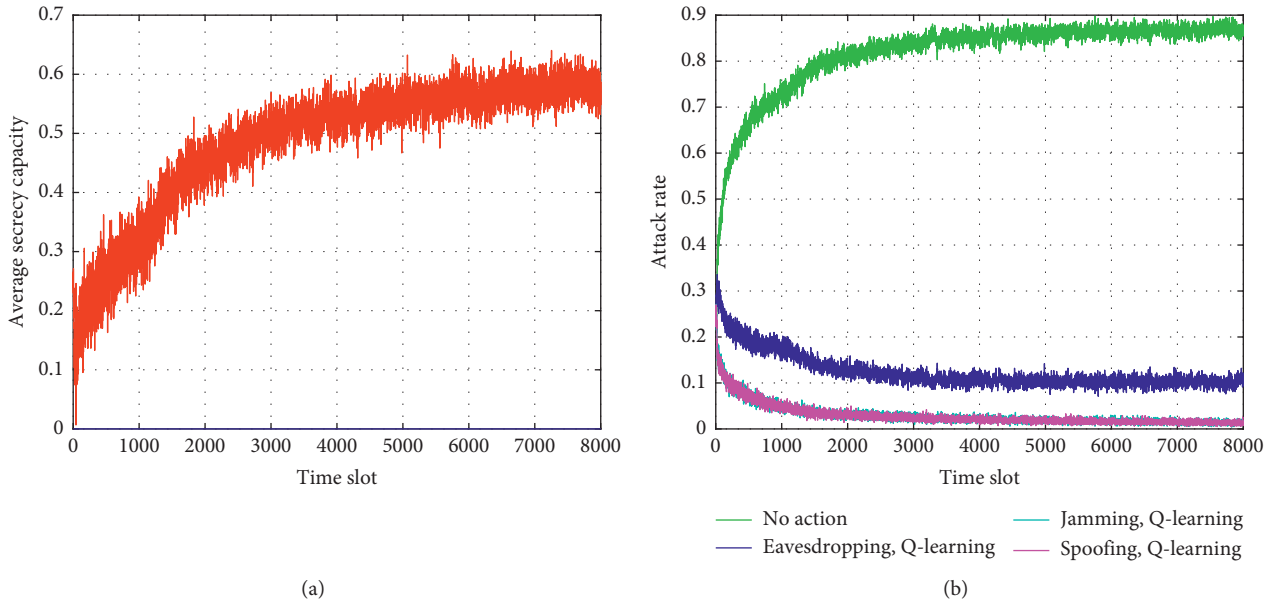


Figure 4: Continued.

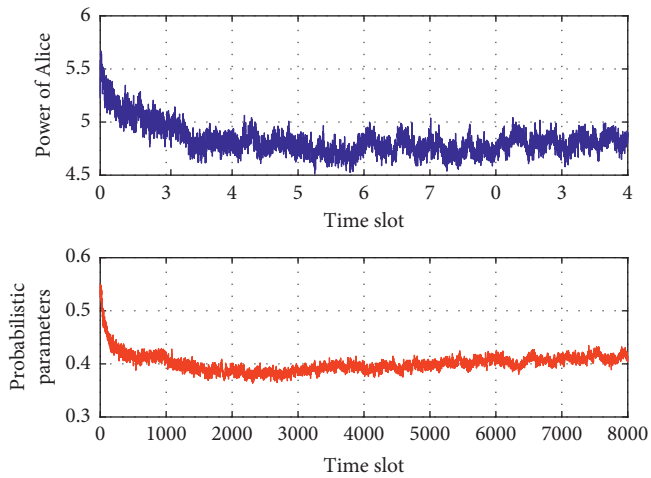


(c)

Figure 4: Illustration of the *NE* in the eavesdropping mode by employing the power control strategy only in the eavesdropping system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power P.



(a) (b)



(c)

Figure 5: Illustration of the *NE* in the no-attack mode by employing both the power control strategy and the probabilistic transmission scheme in the eavesdropping system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit powers P and P_T.

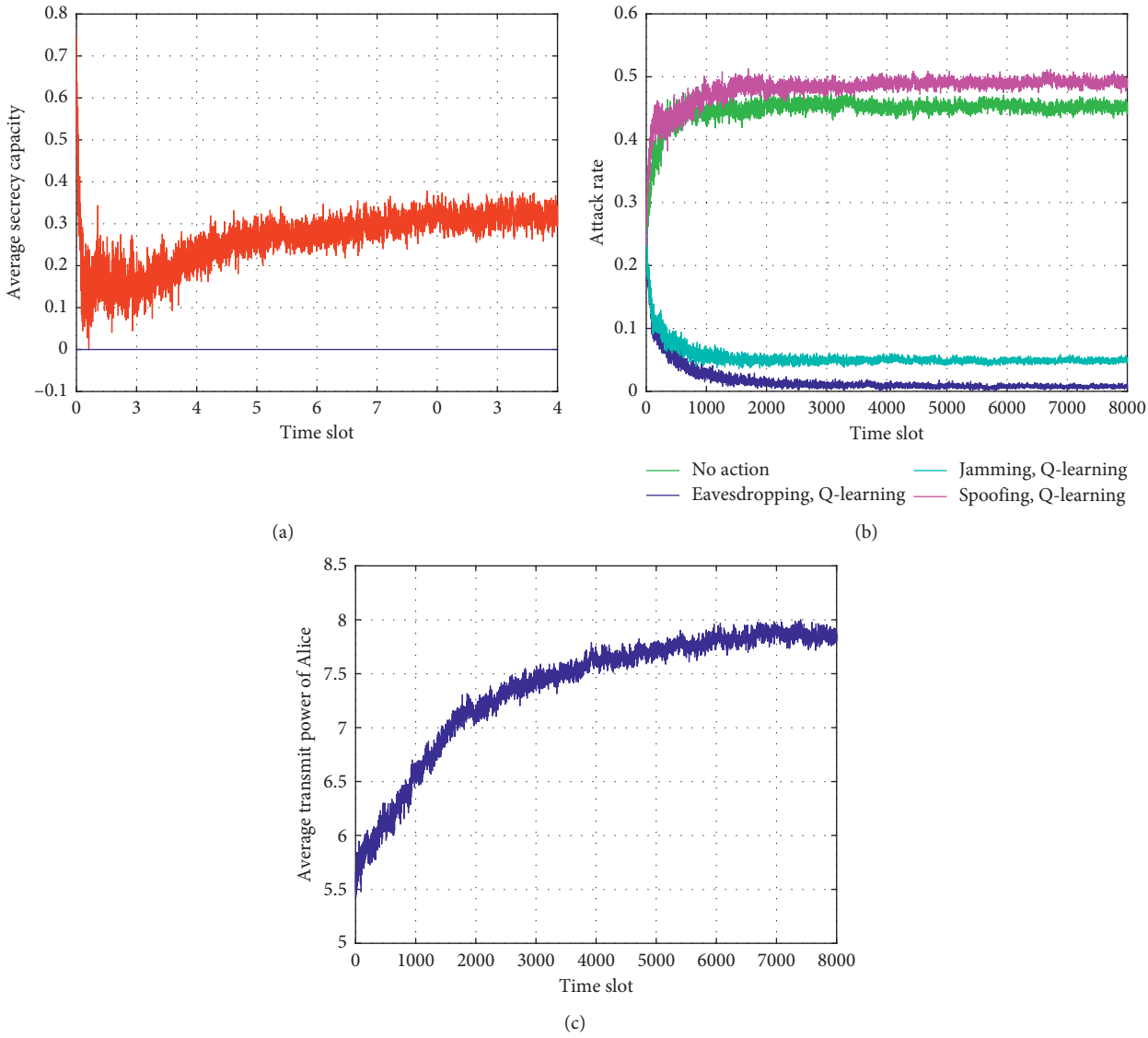


FIGURE 6: Illustration of the *NE* in the spoofing mode by employing the power control strategy only in the spoofing system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power *P*.

secure transmission from *Alice* to *Bob* can be realized, as illustrated in Figure 5(a). Since the *Alice-Eve* link is noticeably better than the *Alice-Bob* link, there is some expected loss in the realized secrecy capacity, when compared with the realized secrecy capacity in the benchmark system, in which the same *Alice-Bob* link is assumed. From Figure 5(c), we may see that the loss in the realized secrecy capacity can be explicated by the low transmit power at *Alice* and the low transmit probability, which is the result of the game decision when the proposed probabilistic transmission policy is utilized.

5.3. Ability to Suppress the Spoofing. Let us consider the following channel conditions as a typical spoofing setup: $\{\lambda_1^{ba} = 2.77, \lambda_2^{ba} = 1.81, \lambda_3^{ba} = 1.07\}$, $\{\lambda_1^{ea} = 1.15, \lambda_2^{ea} = 0.75, \lambda_3^{ea} = 0.44\}$, and $\{\lambda_1^{be} = 9.03, \lambda_2^{be} = 5.16, \lambda_3^{be} = 2.10\}$. Compared with the benchmark system in Section 5.1 and the eavesdropping system in Section 5.2, now the *Eve-Bob* link is the best among the three involved links, the *Alice-Eve* link is the worst, and the *Alice-Bob* channel remains unchanged. According to our analysis in Section 3.2, now the traditional game decision will incline to the spoofing by *Eve*. As illustrated in Figure 6(b), now the spoofing probability by *Eve* is about 50%. In order to resist the spoofing attack, *Alice* needs a relatively high transmit power to achieve the relatively small secrecy capacity, as illustrated in Figures 6(a) and 6(c). Here, we can clearly see that the power control strategy only cannot effectively resist the *Eve*'s motivation to spoof the reception at *Bob*.

In the same spoofing setting, if the proposed adaptive probabilistic transmission scheme is utilized along with the power control strategy, the realized secrecy rate, different attack mode probabilities by *Eve*, the required transmit power, and the probabilistic transmission control at *Alice* are illustrated in Figure 7. We can see that the game results now

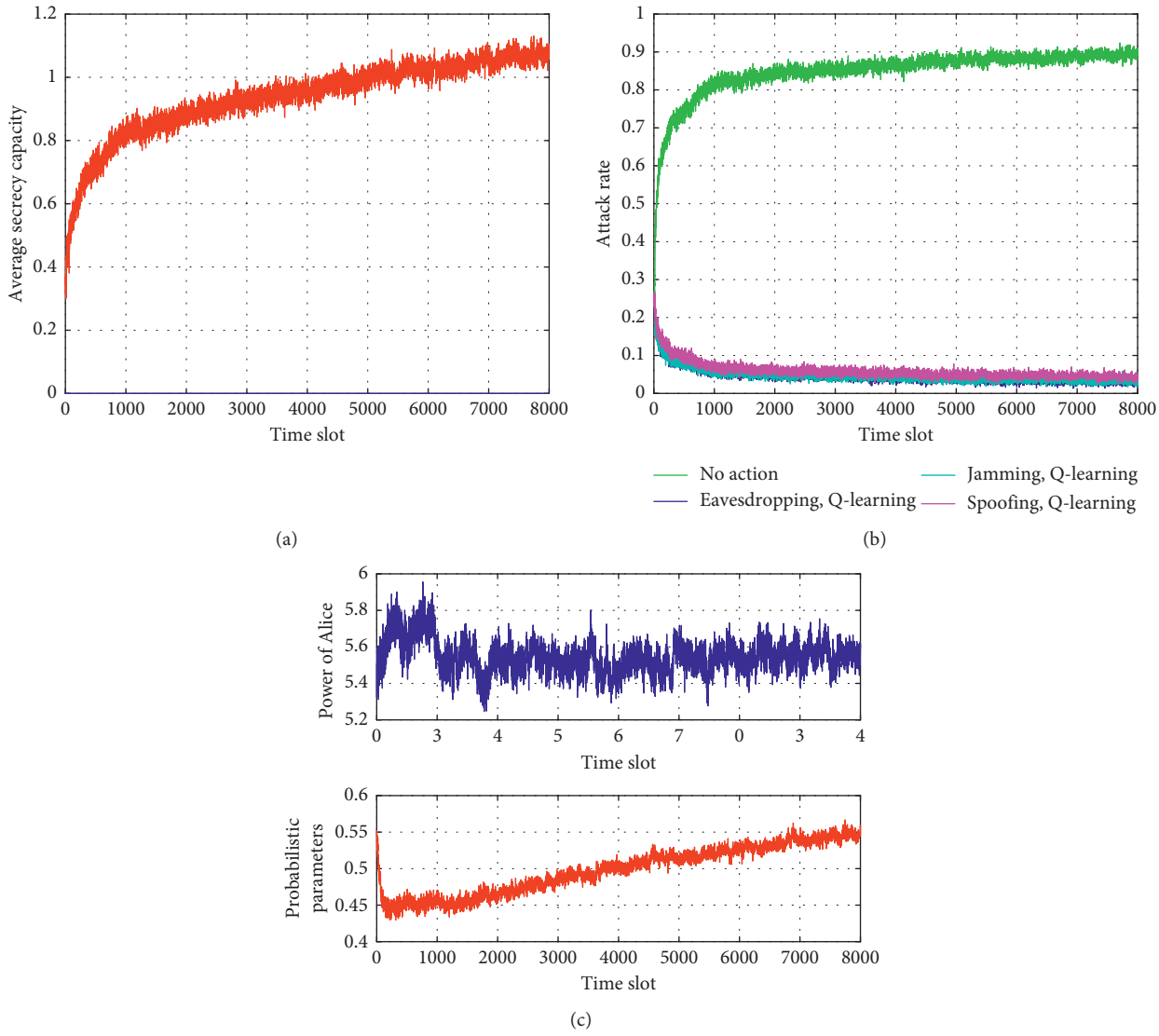


FIGURE 7: Illustration of the *NE* in the no-attack mode by employing both the power control strategy and the probabilistic transmission scheme in the spoofing system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power P and P_T.

incline to *Alice* again. As illustrated in Figure 7(b), now the no-attack probability is about 90%. One may also note that by lowering the transmission probability of *Alice*, we can effectively suppress the attack motivation by *Eve*. As a result, improved secure transmission from *Alice* to *Bob* can be realized, as illustrated in Figure 7(a). Compared with the required transmit power in Figures 6(c) and 7(c), one may readily observe that, with the proposed probabilistic transmission control mechanism, less transmit power is needed to counteract the spoofing attack by *Eve* to realize a better secrecy capacity, which is obviously attractive for practical applications.

5.4. Ability to Suppress the Jamming. Let us consider the following channel conditions as a typical jamming setup: $\{\lambda_1^{ba} = 7.39, \lambda_2^{ba} = 4.84, \lambda_3^{ba} = 2.87\}$, $\{\lambda_1^{ea} = 1.84, \lambda_2^{ea} = 1.21, \lambda_3^{ea} = 0.71\}$, and $\{\lambda_1^{be} = 8.43, \lambda_2^{be} = 4.81, \lambda_3^{be} = 1.96\}$. Compared with the benchmark system in Section 5.1, the eavesdropping system in Section 5.2, and the spoofing system in Section 5.3, now we have a much better *Alice-Bob* link, the *Eve-Bob* link is in good conditions, and the *Alice-Eve* link is the worst. According to our analysis in Section 3.2, now the traditional game decision will incline to the jamming by *Eve*. As illustrated in Figure 8(b), now the jamming probability by *Eve* is about 50%. *Alice* needs a relatively high transmit power to resist the jamming attack, as illustrated in Figure 8(c). Here, we can clearly see that the power control strategy only cannot effectively suppress the *Eve*'s motivation to obstruct the reception at *Bob*.

In the same jamming setting, if the proposed adaptive probabilistic transmission scheme is utilized along with the power control strategy, the realized secrecy rate, different attack mode probabilities by *Eve*, the required transmit power, and the probabilistic transmission control at *Alice* are illustrated in Figure 9. We can see that the game results now

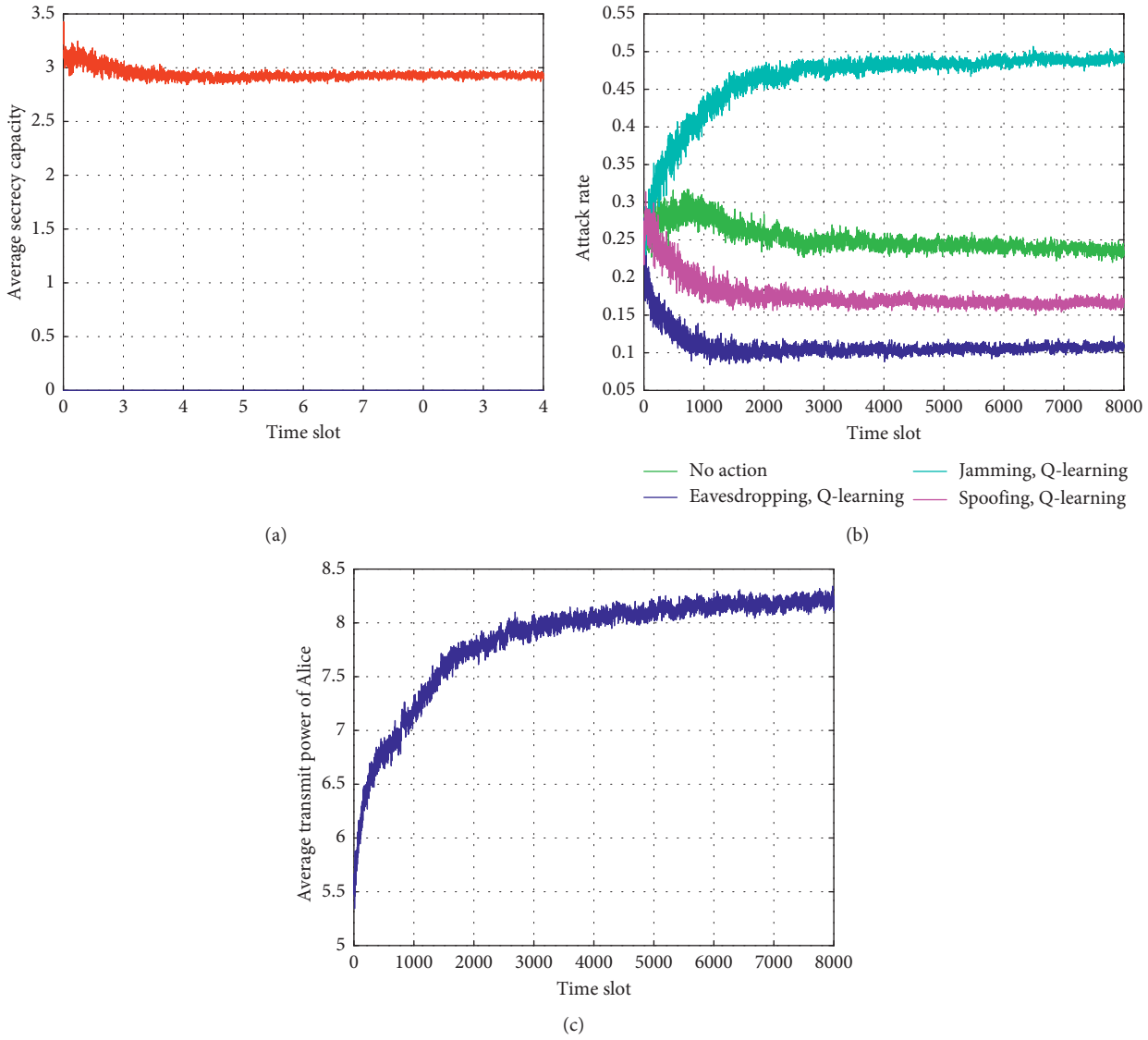


FIGURE 8: Illustration of the *NE* in the jamming mode by employing the power control strategy only in the jamming system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power *P*.

incline to *Alice* again. As illustrated in Figure 9(b), now the no-attack probability is about 50%. One may also note that, by lowering the transmission probability of *Alice*, we can effectively suppress the attack motivation by *Eve*. As a result, improved secure transmission from *Alice* to *Bob* can be realized, as illustrated in Figure 9(a). Compared with the required transmit power in Figures 9(c) and 8(c), one may readily observe that, with the proposed probabilistic transmission control mechanism, less transmit power is needed to counteract the jamming attack by *Eve* to realize a better secrecy capacity, which is desired in practical applications.

In summary, our analysis results confirm us that the proposed adaptive probabilistic MIMO transmission scheme do provide us an effective method to make the MIMO transmitter better resist malicious attackers in adverse channel conditions, in which not only the transmit power of the MIMO transmitter but also the transmission probability will be adjusted to suppress the attack motivation. Meanwhile, we can also conclude from the four typical settings that (1) when the adaptive transmit power policy is enough to suppress the attack motivation, the use of both the adaptive transmit power and the probabilistic transmission control will lead to a more energy efficient MIMO transmission scheme, but with some loss in the realized secrecy capacity; (2) when the adaptive transmit power policy is no longer enough to suppress the attack motivation, the use of both the adaptive transmit power and the probabilistic transmission control will be highly recommended, not only in terms of the attack motivation suppression but also from the perspective of the realized secrecy capacity and the improved energy efficiency. In terms of the realized secrecy capacity, one may also note that the proposed adaptive probabilistic transmission control policy seems to be very attractive in spoofing and jamming scenarios.

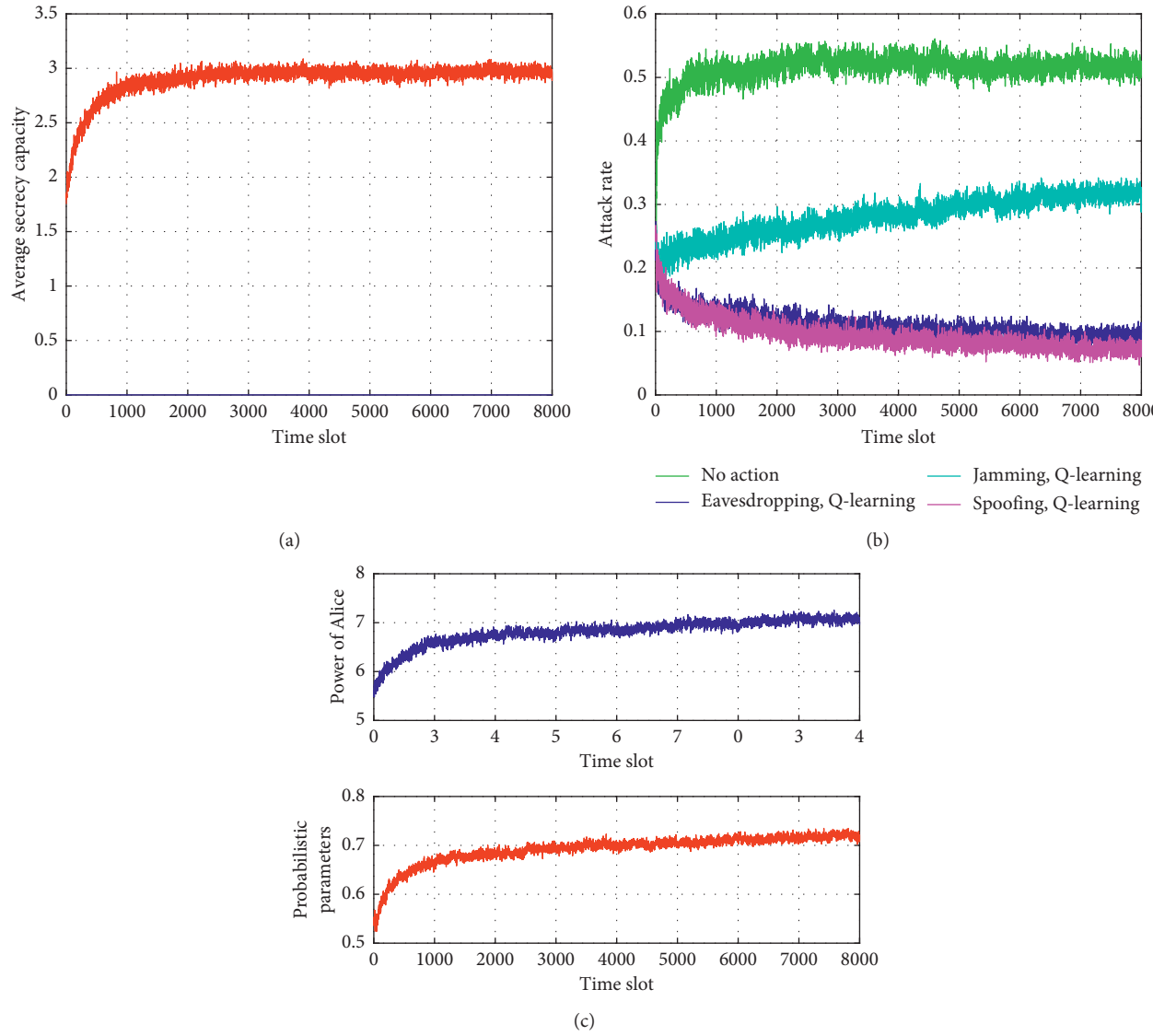


FIGURE 9: Illustration of the *NE* in the no-attack mode by employing both the power control strategy and the probabilistic transmission scheme in the jamming system. (a) Secrecy capacity. (b) Attack probability by *Eve*. (c) Transmit power P and P_T.

6. Conclusion

In this paper, we focus on the noncollaborative game between an MIMO transmitter and one smart malicious attacker, both of which try to maximize their predefined utilities. By carefully analyzing the Nash Equilibrium (NE) and the critical conditions that affect the achieved NE, an adaptive probabilistic MIMO transmission scheme was proposed to make the MIMO transmitter better resist the malicious attacker in adverse channel conditions. Compared with the existing game-based strategy, not only the transmit power of the MIMO transmitter but also the transmission probability will be tuned in the proposed adaptive probabilistic transmission scheme. And our analysis results unveil that the proposed adaptive probabilistic transmission can be regarded as a generalized version of the previous adaptive transmission scheme, which can significantly suppress the attack motivation by the smart attacker and improve the secrecy capacity, even if the adaptive power control strategy fails. Other sophisticated strategies can also be employed to further improve the adaptive secure MIMO transmission scheme. For instance, just like [20], when full-duplex (FD) *Bob* is assumed, some subsets of antenna at *Bob* can be utilized to send the artificial noise to obstruct the eavesdropping by *Eve*. We leave this in the next step work.

Data Availability

In our work, all the data are generated by the simulation platform developed by ourselves, instead of any other data set. Of course, in order to make sure that our data set is properly generated, we have verified the results by ensuring all the results comply with the existing publications, for instance, [18] and [2].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant no. 61771406 and the Key Research Program on Innovation in Industry, University and Research by Guangzhou University. The work of Q. Chen was also supported by Lingnan Yingjie Project of Guangzhou Municipal Government.

References

[1] Y. Tung, S. Han, D. Chen, and K. Shin, "Vulnerability and protection of channel state information in multiuser MIMO networks," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*, pp. 775–786, Scottsdale, AZ, USA, November 2014.

[2] 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.

[3] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2564–2573, May 2012.

[4] A. Mukherjee and A. L. Swindlehurst, "Optimal strategies for countering dual-threat jamming/eavesdropping-capable adversaries in MIMO channels," in *Proceedings of the IEEE Military Communications Conference (MILICOM)*, pp. 1695–1700, San Jose, CA, USA, November 2010.

[5] X. He, H. Dai, P. Ning, and R. Dutta, "A stochastic multi-channel spectrum access game with incomplete information," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 4799–4804, London, UK, June 2015.

[6] A. Garnaev, M. Baykal-Gursoy, and H. V. Poor, "Incorporating attack-type uncertainty into network protection," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 8, pp. 1278–1287, 2014.

[7] A. Mukherjee and A. L. Swindlehurst, "Jamming games in the MIMO wiretap channel with an active eavesdropper," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 82–91, 2013.

[8] A. Garnaev and W. Trappe, "The eavesdropping and jamming dilemma in multi-channel communications," in *Proceedings of the 2013 IEEE International Conference on Communications (ICC)*, pp. 2160–2164, Budapest, Hungary, June 2013.

[9] 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, 2016.

[10] C. Xie and L. Xiao, "User-centric view of smart attacks in wireless networks," in *Proceedings of the of 2016 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)*, Nanjing, China, October 2016.

[11] L. Xiao, G. Sheng, X. Wan, W. Su, and P. Cheng, "Learning-based PHY-layer authentication for underwater sensor networks," *IEEE Communications Letters*, vol. 23, no. 1, pp. 60–63, 2019.

[12] 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.

[13] A. Rahmati and H. Dai, "Reinforcement learning for interference avoidance game in RF-powered backscatter communications," in *Proceedings of the of 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019.

[14] 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, 2018.

[15] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, USA, 1998.

[16] E. R. Gomes and R. Kowalczyk, "Dynamic analysis of multiagent Q-learning with ε-greedy exploration," in *Proceedings of the ACM 26th Annual International Conference on Machine Learning (ICML 09)*, pp. 369–376, Montreal, Canada, June 2009.

[17] M. Wunder, M. Littman, and M. Babes, "Classes of multiagent Q-learning dynamics with ε-greedy exploration," in *Proceedings of the ACM Annual International Conference Machine Learning (ICML)*, pp. 1167–1174, Haifa, Israel, 2010.

[18] Y. Li, L. Xiao, H. Dai, and H. V. Poor, "Game theoretic study of protecting MIMO transmissions against smart attacks," in *Proceedings of the IEEE International Conference on Communications*, London, UK, May 2017.

[19] A. Goldsmith, *Wireless Communications*, Cambridge University Press, Cambridge, UK, 2005.

[20] L. Li, A. P. Petropulu, and Z. Chen, "MIMO secret communications against an active eavesdropper," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2387–2401, 2017.