UAV Relay in VANETs Against Smart Jamming with Reinforcement Learning

Liang Xiao, *Senior Member, IEEE,* Xiaozhen Lu, Dongjin Xu, Yuliang Tang, *Member, IEEE,* Lei Wang, *Member, IEEE* and Weihua Zhuang, *Fellow, IEEE*

***Abstract*—Frequency hopping-based anti-jamming techniques are not always applicable in vehicular ad-hoc networks (VANETs) due to the high mobility of onboard units (OBUs) and the large-scale network topology. In this paper, we use unmanned aerial vehicles (UAVs) to relay the message of an OBU and improve the communication performance of VANETs against smart jammers that observe the ongoing OBU and UAV communication status and even induce the UAV to use a specific relay strategy and then attack it accordingly. More specifically, the UAV relays the OBU message to another roadside unit (RSU) with a better radio transmission condition if the serving RSU is heavily jammed or interfered. The interactions between a UAV and a smart jammer are formulated as an anti-jamming UAV relay game, in which the UAV decides whether or not to relay the OBU message to another RSU, and the jammer observes the UAV and the VANET strategy and chooses the jamming power accordingly. The Nash equilibria of the UAV relay game are derived to reveal how the optimal UAV relay strategy depends on the transmit cost and the UAV channel model. A hotbooting policy hill climbing (PHC)-based UAV relay strategy is proposed to help the VANET resist jamming in the dynamic game without being aware of the VANET model and the jamming model. Simulation results show that the proposed relay strategy can efficiently reduce the bit error rate of the OBU message and thus increase the utility of the VANET compared with a Q-learning based scheme.**

***Index Terms*—VANET, jamming, reinforcement learning, game theory, unmanned aerial vehicles.**

I. Introduction

Vehicular ad-hoc networks (VANETs) support vehicle-to-vehicle communications and vehicle-to-infrastructure communications to improve the transmission security, help build unmanned-driving, and support booming applications of on-board units (OBUs) [1]. The high mobility of OBUs and the large-scale dynamic network with fixed roadside units (RSUs) make the VANET vulnerable to jamming [2]. A jammer sends faked or replayed signals and aims to block the ongoing transmissions between OBUs and the serving RSUs. By applying smart radio devices to observe the ongoing VANET communication and evaluate the underlying policy, a smart jammer not only has flexible control over the jamming frequencies and signal strengths but also induces the VANET to use a specific communication strategy and then attacks it accordingly.

The anti-jamming communication of VANETs can be significantly improved by using unmanned aerial vehicles (UAVs) to relay the OBU message. Being faster to deploy, UAVs generally have better channel states due to the line-of-sight (LOS) links and smaller path-loss exponents [3], [4] when they communicate with OBUs and RSUs, compared with the serving RSUs at a fixed location on the ground that might be severely blocked by a smart jammer. Therefore, UAVs help relay the OBU message to improve the signal-to-interference-plus-noise-ratio (SINR) of the OBU signals, and thus reduce the bit-error-rate (BER) of the OBU message, especially if the serving RSUs are blocked by jammers and/or interference.

Game theory has been used to study the anti-jamming power control in wireless networks [5], [6]. However, to our best knowledge, the jamming resistance in the UAV-aided VANETs is still an open problem. In this paper, we formulate an anti-jamming UAV relay game, in which the UAV decides whether or not to relay an OBU message to another RSU, and the smart jammer chooses its jamming power under a UAV channel model with the path loss, log-normal shadowing and Rayleigh fading [7]. The Nash equilibria (NEs) of the game are derived and the conditions that each NE exists are provided to disclose how the VANET transmission, the jamming model and the UAV channel model impact the BER of the OBU message in the UAV-aided VANET against jamming.

The UAV relay process in the dynamic game can be viewed as a Markov decision process (MDP), and thus reinforcement learning (RL) such as Q-learning can achieve the optimal strategy via trials-and-errors if the game is long enough [8], [9]. As a model-free RL technique for the mixed-strategy game, the policy hill climbing (PHC) algorithm can achieve the optimal policy without knowing the jamming model and the VANET model [10]. Compared with Q-learning, the PHC-

This paper was supported in part by National Natural Science Foundation of China under Grant 61671396, 61671253 and 91638204, in part by the open research fund of National Mobile Communications Research Laboratory, Southeast University No.2018D08, in part by the Open Research Fund of the State Key Laboratory of Integrated Services Networks, Xidian University ISN17-04, and in part by the Major Projects of the Natural Science Foundation of the Jiangsu Higher Education Institutions 16KJA510004.

Liang Xiao is with the Department of Communication Engineering, Xiamen University, Xiamen 361005, China, and also with the National Mobile Communications Research Laboratory, Southeast University, China. E-mail: lxiao@xmu.edu.cn.

Xiaozhen Lu, Dongjin Xu and Yuliang Tang are with the Department of Communication Engineering, Xiamen University, Xiamen 361005, China.

Lei Wang is with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China. E-mail: wanglei@njupt.edu.cn.

Weihua Zhuang is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1. E-mail: wzhuang@uwaterloo.ca.

based relay strategy that provides more randomness in the decision can fool the jammers with uncertainty. Therefore, we propose the PHC-based UAV relay strategy to improve the anti-jamming performance of the VANET communication. In this scheme, the jammer cannot predict the optimal relay policy and thus cannot attack the UAV accordingly.

As a type of transfer learning [11], the hotbooting technique is applied to accelerate the learning process and thus improve the VANET communication performance against jamming. More specifically, the hotbooting technique initializes the value of the quality function or Q-function for each action-state pair with the experiences in similar scenarios to avoid the random exploration of the standard PHC that initializes the Q-function with an all-zero matrix. Simulation results show that our proposed hotbooting PHC-based relay strategy can efficiently reduce the BER of the OBU message and increase the utility of the UAV.

The contributions of this work can be summarized as follows:

- We formulate a UAV relay game to study the resistance against smart jamming in the UAV-aided VANET. The NEs of the game and the conditions under which the NEs exist are provided.
- We propose a hotbooting PHC-based UAV relay strategy to resist smart jamming without the knowledge of the UAV channel model and the jamming model. Simulation results show that this scheme achieves a lower BER of the OBU message and a higher utility compared with a Q-learning based relay strategy.

The rest of this paper is organized as follows. We review the related work in Section II and present the system model of the UAV-aided VANET in Section III. We present the anti-jamming UAV relay game in Section IV and study the anti-jamming UAV relay stochastic game in Section V. A hotbooting PHC-based UAV relay strategy is presented in Section VI. Simulation results are provided in Section VII and conclusion is drawn for this work in Section VIII.

II. Related Work

The distributed medium access control (MAC) scheme as proposed in [12], [13] supports the emergency message dissemination to reduce the transmission delay in VANETs. The vehicle-assisted data delivery protocols as proposed in [14] can forward a packet to a new vehicle on the road with the shortest data-delivery delay in VANETs. The jamming detection model formulated in [15] provides the correlation between jamming detection error and data reception time of the VANETs. The jamming defense game as formulated in [16] helps design the defense resource allocation in VANETs.

UAVs have been used to relay mobile messages for ground terminals. The optimization of multi-antenna UAVs and mobile ground terminals in [7] improves the uplink sum rate in a wireless relay network. The relay scheme as proposed in [17] uses UAVs to maximize the capacity of a wireless relay network. The UAV-aided intrusion detection scheme presented in [3] uses UAVs to relay the alarm messages regarding lethal attacks of vehicles to improve the detection accuracy and reduce the energy consumption in vehicular networks. The iterative UAV relay algorithm proposed in [18] optimizes the power control and relay trajectory to improve the throughput of mobile networks. The UAV-enabled mobile relaying system as investigated in [19] uses the difference-of-concave program to optimize the transmit power of the mobile device and the relay, and maximize the secrecy data rate.

The UAV placement strategy as developed in [20] enhances the coverage of public safety communications. The UAV-aided sensor deployment in [21] improves the localization and navigation to monitor post-disaster areas. The field tests in [22] show the impact of the UAV altitude on the communication quality in autonomous vehicles. The UAV-aided wireless sensor network as investigated in [23] can reduce the packet loss and power consumption of the network against node failures. The UAV-assisted data gathering system as developed in [24] reduces the required execution time and the energy consumption in wireless sensor networks. The tradeoff between the coverage and the time required to cover the entire target area of the UAV is studied in [25] to determine the number of stop points of the UAV in an underlaid device-to-device communication network.

The ambient noise immunity based anti-jamming algorithm developed in [26] adjusts the false detection threshold to improve the packet delivery rate in VANETs. The hideaway strategy as proposed in [2] determines when to keep silent based on the packet transmission ratio to improve jamming resistance. The jamming detection scheme as proposed in [27] can improve the message invalidation ratio in time-critical networks. The MAC-based jamming detection scheme as presented in [28] reduces the false alarm rate and the time required to monitor vehicular networks.

Reinforcement learning techniques have been widely applied to improve security in wireless networks [8], [9], [29]. The non-cooperative power control algorithm as presented in [29] in the repeated game can improve the throughput of wireless ad hoc networks. The prospect theory based dynamic game as formulated in [8] shows the impact of the subjectivity of end-users and jammers on the throughput of cognitive radio networks with Q-learning algorithm. The deep Q-network algorithm as proposed in [9] uses both frequency and spatial diverting to improve the SINR of the signals and the utility of the secondary user in cognitive radio networks.

In [30], we formulated a UAV-aided VANET transmission game with several OBUs and RSUs. In this paper, we extend the study to the anti-jamming UAV relay stochastic game with random channel power gains for the OBU-UAV link and the OBU-RSU_1 link, and consider the impact of the UAV-RSU_2 distance and the speed of the OBU on the BER of the OBU message. We also evaluate the dynamic games with a smart jammer that changes the jamming policy according to the learning algorithm of the UAV.

III. UAV-aided VANETs

A. Network Model

In this work, we consider an OBU that moves along the road at a speed denoted by $v^{(k)} \in [0, \text{V}]$ at time slot k, where

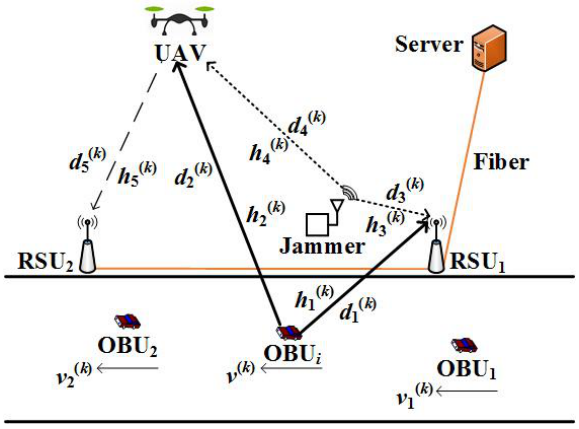


Fig. 1: Illustration of a UAV-aided VANET, in which the OBU moving with speed $v^{(k)}$ sends a message at time slot k to a server via the serving RSU (RSU_1) and the UAV that is connected with RSU_2 and is less affected by the jammer.

V is the maximum speed and time is partitioned into slots of a constant duration. The OBU aims to send a message to a server via several RSUs and a UAV in a time slot, as illustrated in Fig. 1. The RSUs at fixed locations are connected via fibers with each other and the server. Equipped with sensors such as cameras and a global positioning system receiver, the OBU gathers the sensing information and sends a message to the server via the serving RSU denoted by RSU_1. We assume that both the UAV and RSU_1 receive the message from the OBU and then the UAV decides whether to connect to the server via RSU_2 afterwards in the time slot. For simplicity, the constant channel power gains are assumed to be constant in each time slot.

Let $\mathbf{d}^{(k)} = [d_1^{(k)}, d_2^{(k)}, d_3^{(k)}, d_4^{(k)}, d_5^{(k)}]$ denote the topology vector of the network at time slot k, where $d_1^{(k)}$ denotes the distance between the OBU and RSU_1, $d_2^{(k)}$ is the distance between the OBU and the UAV, $d_3^{(k)}$ corresponds to the distance of the jammer-RSU_1 link, $d_4^{(k)}$ is the distance between the jammer and the UAV, and $d_5^{(k)}$ is the distance between the UAV and RSU_2. The distances $d_2^{(k)}$ and $d_1^{(k)}$ depend on the speed of the OBU at time slot k. The OBU sends a message to a server via RSU_1 and the UAV that is connected with RSU_2 at time slot k with a fixed transmit power $P^{(k)}$. The SINR of the signals received by r ($r = 1$ for RSU_1 or $r = 2$ for the UAV) sent from the OBU and the SINR of the signals received by RSU_2 sent from the UAV at time slot k are denoted by $\rho_r^{(k)}$ and $\rho_3^{(k)}$, respectively. The BER of a signal denoted by $p_e(\rho)$ depends on the SINR per bit, ρ, of the received signal for a given modulation mode. The BER of the OBU message at time slot k denoted by $P_e^{(k)}$ depends on the OBU-RSU_1 link, the OBU-UAV link and the UAV-RSU_2 link, and is given by

$$P_e^{(k)} = \min\left(p_e\left(\rho_1^{(k)}\right), p_e\left(\min\left(\rho_2^{(k)}, \rho_3^{(k)}\right)\right)\right). \quad (1)$$

The BER of the message depends on the minimum of the BER of the OBU-RSU_1 signal as shown in the first term, and the BER of the weaker signal of the OBU-UAV link and the UAV-RSU_2 link at that time as represented in the second term. According to the channel quality and the BER of the OBU message, the UAV decides whether or not to relay the OBU message to RSU_2, which is denoted by $x \in \mathbf{A} = \{0, 1\}$, where $\mathbf{A}$ is the feasible action set of the UAV. The UAV relays the OBU message at time slot k with a fixed transmit power $P_U^{(k)}$ and a transmit energy cost $C_U^{(k)}$ if $x = 1$, and keeps silent otherwise. The system server can gather the sensing message sent from the OBU via RSU_1 or the UAV.

The jammer at a fixed location sends faked or replayed signals at time slot k to block RSU_1 with a jamming cost $C_J^{(k)}$ estimates the transmit power of the OBU and the UAV. In addition, the smart jammer applies smart radio devices to eavesdrop the control channel of the VANET and estimate the VANET transmission policy. According to the estimated VANET communication, the smart jammer changes its jamming power $y \in [0, P_J^M]$, where P_J^M is the maximum jamming power. For simplicity, we assume a constant noise power in the received signal denoted by σ and the jammer is too far away from the UAV and RSU_2 to block them.

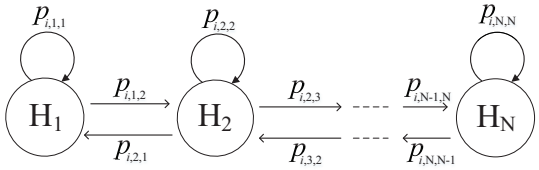


Fig. 2: Markov chain based channel model between the OBU and RSU_1 with N states, where $p_{i,m,n}$ is the probability that h_i changes from H_m to H_n in a time slot.

B. Channel Model

The channel power gain vector of the system denoted by $\mathbf{h}^{(k)} = [h_1^{(k)}, h_2^{(k)}, h_3^{(k)}, h_4^{(k)}, h_5^{(k)}]$ at time slot k consists of the channel power gain of the OBU-RSU_1 link $h_1^{(k)}$, the OBU-UAV link $h_2^{(k)}$, the jammer-RSU_1 link $h_3^{(k)}$, the jammer-UAV link $h_4^{(k)}$, and the UAV-RSU_2 link $h_5^{(k)}$. Similar to [19], the channel gain is modeled as

$$h_i^{(k)} = \theta_0 \Delta_i^{(k)} \left(\frac{d_i^{(k)}}{d_0}\right)^{-\alpha_i} \quad (2)$$

where θ_0 is the channel power gain at the reference distance d_0, the channel time variation $\Delta_i^{(k)}$ depends on the Doppler shift due to the node mobility. The path-loss exponent α_i is set according to [31] and [32], e.g., $\alpha_i = 2$ for the OBU-UAV, jammer-UAV and UAV-RSU_2 links and $\alpha_i = 4$ otherwise. The path loss of the jammer-UAV radio link is assumed to be much higher than that of the jammer-RSU_1 link due to a longer distance. Similarly, the radio link between the OBU and the UAV has a higher path loss compared with that of the OBU-RSU_1 link.

Due to the uncorrelated locations of RSU_1 and the UAV, $h_1^{(k)}$ is independent with $h_2^{(k)}$. The UAV transmission fails

TABLE I: Summary of symbols and notations

$P^{(k)}$	Transmit power of the OBU at time slot k
$P_U^{(k)}$	Transmit power of the UAV at time slot k
V	Maximum OBU speed
$v^{(k)}$	Speed of the OBU at time slot k
σ	Received noise power
$\mathbf{h}^{(k)}$	Channel gain vector at time slot k
$\mathbf{d}^{(k)}$	Distance vector at time slot k
$\rho_r^{(k)}$	SINR of the signals received by r from the OBU at time slot k
$\rho_3^{(k)}$	SINR of the signals received by RSU_2 from the UAV at time slot k
$\mathbf{A}$	Action set of the UAV
α	Learning rate of the UAV
δ	Learning discount factor of the UAV
$u_{U/J}^{(k)}$	Utility of the UAV/jammer at time slot k
$U_{U/J}^{(k)}$	Expected utility of the UAV/jammer at time slot k
$C_{U/J}^{(k)}$	Transmit cost of the UAV/jammer at time slot k

if the UAV is too far away from the OBU, and the UAV can cover the whole geographic area if it is high enough. For simplicity, the channel power gain is quantized into N levels with $h_i^{(k)} \in \{H_a\}_{1\le a\le N}, i = 1, 2$, and is modeled as a Markov chain with N states. As shown in Fig. 2, the transition probability of the channel gain h_i from H_m to H_n during time slot k denoted by $p_{i,m,n}^{(k)}$ depends on the OBU speed given by

$$p_{i,m,n}^{(k)} = \Pr\left(h_i^{(k)} = H_n \mid h_i^{(k-1)} = H_m\right). \tag{3}$$

For ease of reference, some important notations are summarized in TABLE 1.

IV. Anti-jamming Transmission Game

The interactions between the UAV and the jammer are formulated as an anti-jamming relay game. In this game, the UAV decides whether or not to relay the OBU message to RSU_2, $x \in \mathbf{A}$, while the smart jammer chooses its jamming power $y \in [0, P_J^M]$. The UAV relay decision depends on the channel quality and the BER. The utility of the UAV at time slot k denoted by $u_U^{(k)}$ is based on the SINR of the signal received by the RSUs and the UAV and the transmit cost, which is given by

$$u_U^{(k)}(x,y) = x\max\left(\frac{P^{(k)}h_1^{(k)}}{\sigma+yh_3^{(k)}}, \min\left(\frac{P^{(k)}h_2^{(k)}}{\sigma+yh_4^{(k)}}, \frac{P_U^{(k)}h_5^{(k)}}{\sigma}\right)\right) - xC_U^{(k)} + \frac{(1-x)P^{(k)}h_1^{(k)}}{\sigma+yh_3^{(k)}}. \tag{4}$$

In (4), the first term in the max function represents the SINR of the signal sent by the OBU and received by RSU_1 at time slot k, and the second term represents the SINR of the weaker signal over the OBU-UAV link and the UAV-RSU_2 link at time slot k.

The utility of the jammer at time slot k, denoted by $u_J^{(k)}$, depends on the energy consumption of the UAV and the jamming cost, and is given by

$$u_J^{(k)}(x,y) = -x\max\left(\frac{P^{(k)}h_1^{(k)}}{\sigma+yh_3^{(k)}}, \min\left(\frac{P^{(k)}h_2^{(k)}}{\sigma+yh_4^{(k)}}, \frac{P_U^{(k)}h_5^{(k)}}{\sigma}\right)\right) + xC_U^{(k)} + \frac{(x-1)P^{(k)}h_1^{(k)}}{\sigma+yh_3^{(k)}} - yC_J^{(k)}. \tag{5}$$

The time index k in the superscript is omitted unless necessary.

The NE of the anti-jamming game denoted by (x^*, y^*) corresponds to the best response strategy if the opponent follows the NE strategy. By definition, we have

$$u_U(x^*,y^*) \ge u_U(x,y^*),\ \forall x \in \{0,1\} \tag{6}$$

$$u_J(x^*,y^*) \ge u_J(x^*,y),\ \forall y \in \left[0, P_J^M\right]. \tag{7}$$

Theorem 1. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $(0,0)$*, if*

$$\sigma^2 > \min\left(\frac{Ph_1h_3}{C_J}, \min\left(\frac{(P_Uh_5-Ph_1)^2}{C_U^2}, \frac{P^2(h_2-h_1)^2}{C_U^2}\right)\right) \tag{8}$$

or

$$Ph_1 > \max(P_Uh_5, Ph_2). \tag{9}$$

Proof: By (4), if (8) holds, we have

$$u_U(1,0) = \frac{P_Uh_5}{\sigma} - C_U \le \frac{Ph_1}{\sigma} = u_U(0,0). \tag{10}$$

Thus, (6) holds for $(x^*, y^*) = (0,0)$.

By (5), let $\tilde{y} = \sqrt{Ph_1/(C_Jh_3)} - \sigma/h_3 < 0$, we have

$$u_J(0,y) = -\frac{Ph_1}{\sigma+yh_3} - yC_J \tag{11}$$

$$\frac{\partial u_J(0,\tilde{y})}{\partial y} = \frac{Ph_1h_3}{(\sigma+\tilde{y}h_3)^2} - C_J = 0 \tag{12}$$

and $\partial^2 u_J/\partial y^2 \le 0$, indicating that if (8) holds, $u_J(0,y)$ is concave with respect to (w.r.t.) y. Thus, $\forall y \in [0, P_J^M]$, we have

$$u_J(0,0) = -\frac{Ph_1}{\sigma} \ge -\frac{Ph_1}{\sigma+yh_3} - yC_J = u_J(0,y). \tag{13}$$

Thus, (7) holds for $(x^*, y^*) = (0,0)$, and thus $(0,0)$ is an NE of the game. Similarly, we can prove that $(0,0)$ is an NE of the game if (9) holds. ■

Note that the proof for the following theorems is similar to the proof of Theorem 1 and is omitted. Now we consider the BER of the OBU message in the VANET with quadrature phase-shift keying (QPSK). By (1), we have

$$P_e = \frac{1}{2}\min\left(\text{erfc}\left(\sqrt{\frac{\rho_1}{2}}\right), \text{erfc}\left(\sqrt{\frac{\min(\rho_2,\rho_3)}{2}}\right)\right). \tag{14}$$

Corollary 1. *If (8) holds, the BER of the OBU message in the UAV-aided VANET with QPSK at the NE of the game is*

given by

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\min(Ph_2, P_U h_5)}{2\sigma}}\right). \quad (15)$$

Corollary 2. *If (9) holds, the BER of the OBU message in the UAV-aided VANET with QPSK at the NE of the game is given by*

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{2\sigma}}\right). \quad (16)$$

Remark: Both the jammer and the UAV should keep silent to reduce the power consumption if the jamming cost and the transmit cost are high as shown in (8) or the link between the OBU and RSU_1 is in a good condition as shown in (9). In the case, the BER of the OBU message decreases with the transmit power of the OBU as shown in (15) and (16).

Theorem 2. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $(1, 0)$*, if*

$$C_U\sigma + Ph_1 \leq P_U h_5 < \frac{Ph_2\sigma}{\sigma + P_J^M h_4} \quad (17)$$

or

$$\frac{C_U\sigma}{h_2 - h_1} \leq P < \min\left(\frac{C_J\sigma^2}{h_1 h_4}, \frac{P_U h_5}{h_2}\right) \quad (18)$$

$$h_4 < h_3. \quad (19)$$

Corollary 3. *If (17) holds, the BER of the OBU message in the UAV-aided VANET with QPSK at the NE of the game is given by*

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{P_U h_5}{2\sigma}}\right). \quad (20)$$

Corollary 4. *If (18) and (19) hold, the BER of the OBU message in the UAV-aided VANET with QPSK at the NE of the game is given by*

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_2}{2\sigma}}\right). \quad (21)$$

Remark: The UAV relays for the OBU to improve its performance if the transmit cost is low and the link between the OBU and the UAV is in a good condition as shown in (17). The jammer keeps silent if the jamming cost is high as shown in (18).

Theorem 3. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $\left(0, \sqrt{Ph_1/(C_J h_3)} - \sigma/h_3\right)$*, if*

$$\frac{Ph_1 h_3}{(\sigma + P_J^M h_3)^2} \leq C_J \leq \min\left(\frac{Ph_1 h_3}{\sigma^2}, \max\left(\frac{P_U^2 h_5^2 h_3}{Ph_1\sigma^2}, \frac{Ph_3(h_2 h_3 - h_1 h_4)^2}{h_1\sigma^2(h_3 - h_4)^2}\right)\right) \quad (22)$$

$$C_U \geq \min\left(\frac{P_U h_5}{\sigma} - \frac{Ph_1}{\sigma + P_J^M h_3}, \frac{P(h_2 - h_1)}{\sigma}\right) \quad (23)$$

or

$$Ph_1 C_J\sigma^2 \geq \max\left(P_U^2 h_5^2 h_3, \frac{P^2 h_2^2 h_3^3 \sigma^2 C_J}{\left(\sigma(h_3 - h_4)\sqrt{C_J} + h_4\sqrt{Ph_1 h_3}\right)^2}\right). \quad (24)$$

Theorem 4. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $\left(1, \sqrt{Ph_2/(C_J h_4)} - \sigma/h_4\right)$*, if*

$$C_U \leq \frac{Ph_2}{\sigma + P_J^M h_4} - \frac{Ph_1}{\sigma + P_J^M h_3} \quad (25)$$

$$\max\left(\frac{Ph_1 h_4}{(\sigma + P_J^M h_4)^2}, \frac{Ph_4(h_1 h_4 - h_2 h_3)^2}{h_2\sigma^2(h_4 - h_3)^2}\right) \leq C_J \leq \min\left(\frac{Ph_1 h_4}{\sigma^2}, \frac{P_U^2 h_5^2 h_4}{Ph_2\sigma^2}\right). \quad (26)$$

Theorem 5. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $(0, P_J^M)$*, if*

$$C_J\left(\sigma + P_J^M h_3\right)^2 < Ph_1 h_3 \quad (27)$$

$$C_U \geq \min\left(\frac{P_U h_5}{\sigma} - \frac{Ph_1}{\sigma + P_J^M h_3}, \frac{Ph_2}{\sigma + P_J^M h_4} - \frac{Ph_1}{\sigma + P_J^M h_3}\right) \quad (28)$$

or

$$Ph_1\sigma > \max\left(P_U h_5(\sigma + P_J^M h_3), Ph_2\sigma\right). \quad (29)$$

Theorem 6. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $(1, P_J^M)$*, if*

$$C_U \leq \frac{Ph_2}{\sigma + P_J^M h_4} - \frac{Ph_1}{\sigma + P_J^M h_3} \quad (30)$$

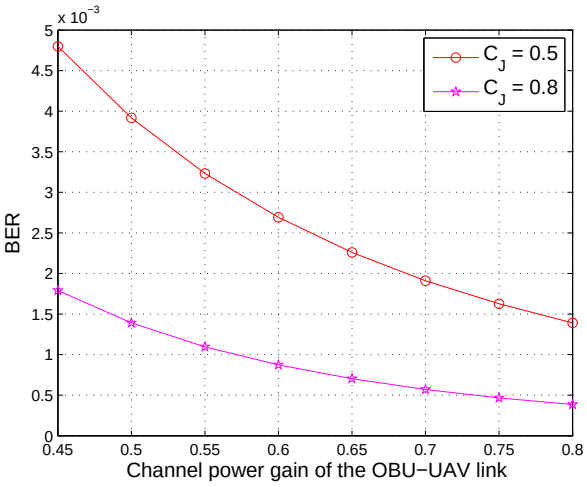
$$\frac{C_J(\sigma + P_J^M h_4)^2}{h_1 h_4} < P \leq \frac{P_U h_5}{h_2} \quad (31)$$

$$h_2\sigma > h_1(\sigma + P_J^M h_4). \quad (32)$$

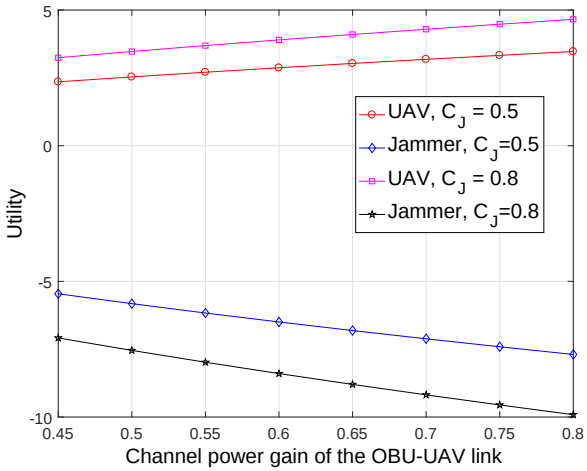
Numerical results with $P = 10$, $P_U = 2$, $\sigma = 0.1$, $h_1 = 0.2$, $h_3 = 0.4$, $h_4 = 0.2$ and $h_5 = 0.3$ in Fig. 3 show that the BER of the OBU message at the NE of the game decreases with the jamming cost C_J. For instance, the BER decreases from 4.8‰ to 1.8‰ and the utility of the UAV increases by 38%, as C_J changes from 0.5 to 0.8, because the smart jammer is less motivated to attack the VANET under a higher jamming cost. The utility of the UAV increases with the channel gain h_2, and the utility of the jammer decreases with it. For example, the utility of the UAV increases by 47.5% as the channel gain h_2 changes from 0.45 to 0.8.

V. Anti-jamming Transmission Stochastic Game

In the anti-jamming relay stochastic game, the UAV decides whether or not relay the OBU message, $x \in \mathbf{A}$, and the smart jammer chooses its jamming power $y \in [0, P_J^M]$ under the time-variant random channel power gains, which can be modeled as a Markov chain with N states. As shown in Fig. 2, the transition probability of h_i at time slot k in the Markov



(a) BER of the OBU message



(b) Utility

Fig. 3: Anti-jamming communication performance of the UAV-aided VANET in the game at the NE, with $P = 10$, $P_U = 2$, $C_U = 1$, $\sigma = 0.1$ and $\mathbf{h} = [0.2, h_2, 0.4, 0.2, 0.3]$.

chain based channel model is denoted by $\mathbf{p}_i^{(k)} = [p_{i,n}^{(k)}]_{1\le n\le N}$, where $p_{i,n}^{(k)}(h_i^{(k-1)}) = \Pr(h_i^{(k)} = H_n \mid h_i^{(k-1)})$.

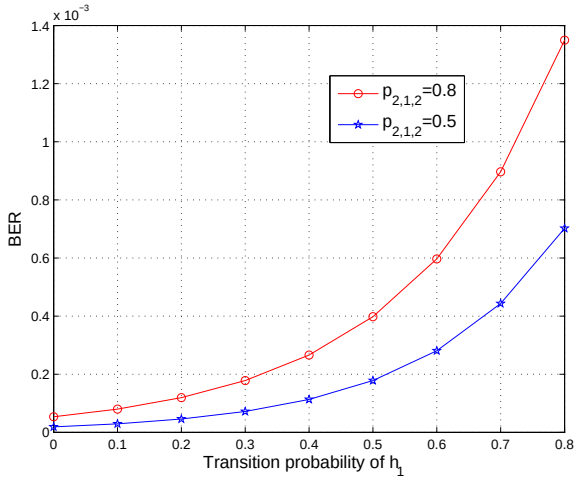
According to the channel model in Fig. 2, the expected utility of the UAV in the stochastic game at time slot k over the $N \times N$ realizations of the channel power gains, $h_1^{(k)}$ and $h_2^{(k)}$, is denoted by $U_U^{(k)}$, and given by (4) as

$$U_U^{(k)}(x,y) = \mathbb{E}_{h_1^{(k)},h_2^{(k)}}\left[u_U^{(k)}(x,y)\right]$$
$$= \sum_{m=1}^{N}\sum_{n=1}^{N} p_{1,m}^{(k)} p_{2,n}^{(k)} x \max\left(\min\left(\frac{P^{(k)}H_n}{\sigma + yh_4^{(k)}}, \frac{P_U^{(k)}h_5^{(k)}}{\sigma}\right), \frac{P^{(k)}H_m}{\sigma + yh_3^{(k)}}\right) + \sum_{m=1}^{N} \frac{(1-x)p_{1,m}^{(k)}P^{(k)}H_m}{\sigma + yh_3^{(k)}} - xC_U^{(k)}. \quad (33)$$

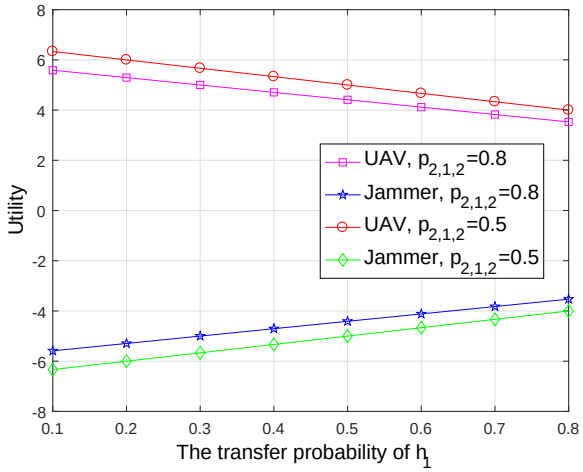
Similarly, the expected utility of the jammer at time slot k in the stochastic game is denoted by $U_J^{(k)}$, and given by (5) as

$$U_J^{(k)}(x,y) = \mathbb{E}_{h_1^{(k)},h_2^{(k)}}\left[u_J^{(k)}(x,y)\right] = xC_U^{(k)} - yC_J^{(k)}$$
$$- \sum_{m=1}^{N}\sum_{n=1}^{N} p_{1,m}^{(k)} p_{2,n}^{(k)} x \max\left(\min\left(\frac{P^{(k)}H_n}{\sigma + yh_4^{(k)}}, \frac{P_U^{(k)}h_5^{(k)}}{\sigma}\right), \frac{P^{(k)}H_m}{\sigma + yh_3^{(k)}}\right) + \sum_{m=1}^{N} \frac{(x-1)p_{1,m}^{(k)}P^{(k)}H_m}{\sigma + yh_3^{(k)}}. \quad (34)$$

Similar to (6) and (7), the NE of the anti-jamming stochastic



(a) BER of the OBU message



(b) Expected utility

Fig. 4: Anti-jamming communication performance of the UAV-aided VANET at the NE, with $P = 10$, $P_U = 2$, $\mathbf{h} = [h_1, h_2, 0.4, 0.2, 0.5]$, $C_U = 1$ and $C_J = 0.8$.

game denoted by (x^*, y^*) is given by definition as follows:

$$U_U(x^*, y^*) \ge U_U(x, y^*), \ \forall x \in \{0, 1\} \quad (35)$$
$$U_J(x^*, y^*) \ge U_J(x^*, y), \ \forall y \in \left[0, P_J^M\right]. \quad (36)$$

Theorem 7. *The anti-jamming transmission stochastic game*

in the UAV-aided VANET has an NE $(0,0)$*, if*

$$C_J\sigma^2 > \sum_{m=1}^{N} p_{1,m} P H_m h_3 \tag{37}$$

$$C_U\sigma \geq \sum_{m=1}^{N}\sum_{n=1}^{N} p_{1,m}p_{2,n} \max\left(PH_m, \min\left(PH_n, P_U h_5\right)\right) - \sum_{m=1}^{N} p_{1,m} P H_m. \tag{38}$$

Proof: By (33), if (38) holds, we have

$$U_U(1,0) = \sum_{m=1}^{N}\sum_{n=1}^{N} p_{1,m}p_{2,n} \max\left(\frac{PH_m}{\sigma}, \min\left(\frac{PH_n}{\sigma}, \frac{P_U h_5}{\sigma}\right)\right) - C_U \leq \sum_{m=1}^{N} \frac{p_{1,m} P H_m}{\sigma} = U_U(0,0). \tag{39}$$

Thus, (35) holds for $(x^*, y^*) = (0,0)$.

By (34), if (37) holds, we have

$$\frac{\partial U_J(0,y)}{\partial y} = \sum_{m=1}^{N} \frac{p_{1,m} P H_m h_3}{(\sigma + y h_3)^2} - C_J < 0 \tag{40}$$

indicating that $U_J(0,y)$ decreases with y, and $\forall y \in [0, P_J^M]$ we have

$$U_J(0,0) = -\sum_{m=1}^{N} \frac{p_{1,m} P H_m}{\sigma} \geq -\sum_{m=1}^{N} \frac{p_{1,m} P H_m}{\sigma + y h_3} - y C_J = U_J(0,y). \tag{41}$$

Similarly, we can prove that (36) also holds for $(x^*, y^*) = (0,0)$, and thus $(0,0)$ is an NE of the game. ■

Algorithm 1 Hotbooting process in the UAV relay

```
Initialize α, β, δ, ξ, s^(0), A
𝔼 = ∅, Q = 0, V = 0, π = 1/A
for n = 1, 2, 3, ..., ξ do
    for k = 1, 2, ..., K do
        Choose x^(k) ∈ A via (47)
        if x^(k) = 1 then
            Relay the OBU message to RSU_2 with a fixed transmit power P_U^(k)
        end if
        Receive the SINR ρ_1, ρ_2, ρ_3, and the BER P_e from server
        Obtain utility u_U^(k)
        Calculate Q*(s^(k), x^(k)) via (44) and (45)
        Calculate π*(s^(k), x^(k)) via (46)
        s^(k+1) = [ρ_1^(k), ρ_2^(k), ρ_3^(k), P_e^(k)]
    end for
end for
Output Q* and π*
```

Remark: Both the jammer and the UAV should keep silent to reduce the power consumption if the jamming cost and the UAV transmit cost are high as shown in (37) and (38).

Theorem 8. *The anti-jamming transmission stochastic game in the UAV-aided VANET has an NE* $(1,0)$*, if*

$$C_J\sigma^2 > \sum_{m=1}^{N} p_{1,m} P H_m h_3 \tag{42}$$

$$C_U\sigma < \sum_{m=1}^{N}\sum_{n=1}^{N} p_{1,m}p_{2,n} \max\left(PH_m, \min\left(PH_n, P_U h_5\right)\right) - \sum_{m=1}^{N} p_{1,m} P H_m. \tag{43}$$

Numerical results with $P = 10$, $P_U = 2$, $\mathbf{h} = [h_1, h_2, 0.4, 0.2, 0.5]$, $C_U = 1$ and $C_J = 0.8$ as shown in Fig. 4 show that the BER of the OBU message at the NE of the stochastic game increases with the transition probability of h_1 and that of h_2. For instance, the BER increases by 72% and the utility of the UAV decreases by 13.3% as $p_{2,1,2}$ changes from 0.5 to 0.8, because the UAV has more trouble to gain the optimal strategies with the unstable channel conditions.

VI. Dynamic Anti-jamming Game with Hotbooting PHC-based Relay Strategy

The repeated interactions between the UAV and the smart jammer in the VANET can be formulated as a dynamic game, in which the jammer determines its jamming power based on the previous VANET transmission, and the UAV chooses its relay strategy based on the system state, which consists of the radio channel state and the BER of the OBU message observed in last time slot. The next system state observed by the UAV is independent of the previous states and actions, for a given system state and UAV relay strategy in the current time slot. Therefore, the UAV relay process in the dynamic game can be viewed as an MDP and the UAV can apply reinforcement learning techniques such as PHC to derive its optimal strategy via trials without the knowledge of jamming model.

In the dynamic game, the UAV decides whether or not to relay the OBU message based on the system state at time slot k denoted by $\mathbf{s}^{(k)}$ that consists of the link quality between the UAV and the OBU, that between RSU_1 and the OBU, the SINR between RSU_2 and the UAV, and the BER of the OBU message at the previous time slot, i.e., $\mathbf{s}^{(k)} = [\rho_1^{(k-1)}, \rho_2^{(k-1)}, \rho_3^{(k-1)}, P_e^{(k-1)}]$.

The learning rate denoted by $\alpha \in (0,1]$ shows the weight of the current experience, and the discount factor $\delta \in [0,1]$ corresponds to the uncertainty on the future utility. The Q-function of the action x at state $\mathbf{s}$ is denoted by $Q(\mathbf{s}, x)$ and is updated in each time slot according to iterative Bellman equation as follows:

$$Q(\mathbf{s}, x) \leftarrow (1-\alpha) Q(\mathbf{s}, x) + \alpha\left(u_U(\mathbf{s}, x) + \delta V(\mathbf{s}')\right), \tag{44}$$

where $\mathbf{s}'$ is the next state if the UAV chooses x at state $\mathbf{s}$, and the value function $V(\mathbf{s})$ maximizes $Q(\mathbf{s}, x)$ over the UAV

Algorithm 2 Hotbooting PHC-based UAV relay strategy

Call Algorithm 1
Initialize α, β, δ, $\mathbf{A}$, $\mathbf{s}^{(0)}$
$\mathbf{Q} = \mathbf{Q}^*$, $\mathbf{V} = \mathbf{0}$, $\pi = \pi^*$
for $k = 1, 2, \ldots$ **do**
  Choose $x^{(k)} \in \mathbf{A}$ via (47)
  if $x^{(k)} = 1$ **then**
    Relay the OBU message to RSU_2 with a fixed transmit power $P_U^{(k)}$
  end if
  Collect the SINR ρ_1, ρ_2, ρ_3, and the BER P_e from server
  Obtain utility $u_U^{(k)}$
  Update $Q\left(\mathbf{s}^{(k)}, x^{(k)}\right)$ via (44)
  Update $V\left(\mathbf{s}^{(k)}\right)$ via (45)
  Update $\pi\left(\mathbf{s}^{(k)}, x^{(k)}\right)$ via (46)
  $\mathbf{s}^{(k+1)} = \left[\rho_1^{(k)}, \rho_2^{(k)}, \rho_3^{(k)}, P_e^{(k)}\right]$
end for

action set given by

$$V(\mathbf{s}) \leftarrow \max_{x \in \{0,1\}} Q(\mathbf{s}, x). \tag{45}$$

The mixed-strategy table in the PHC-based relay denoted by $\pi(\mathbf{s}, x)$ is updated by increasing the probability corresponding to the highest valued action by $\beta \in (0, 1]$, and decreasing other probabilities by $-\beta/(|\mathbf{A}| - 1)$, i.e.,

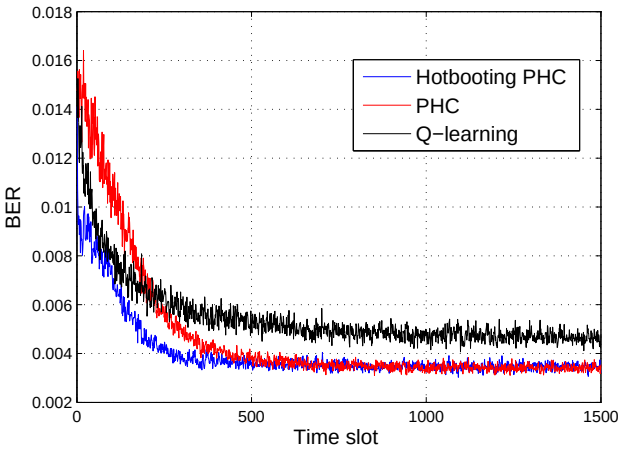
$$\pi(\mathbf{s}, x) \leftarrow \pi(\mathbf{s}, x) + \begin{cases} \beta, & x = \max\limits_{\hat{x} \in \{0,1\}} Q(\mathbf{s}, \hat{x}) \\ \frac{-\beta}{|\mathbf{A}|-1}, & \text{o.w.} \end{cases} \tag{46}$$

The UAV then selects whether or not to relay the OBU message $x \in \mathbf{A}$ according to the mixed strategy $\pi(\mathbf{s}, x)$, i.e.,

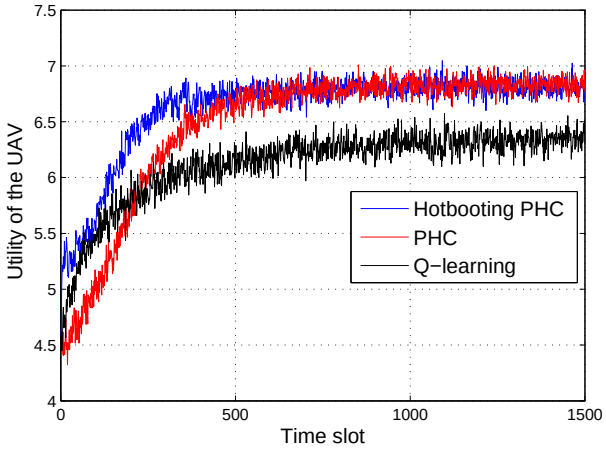
$$\Pr(x = x^*) = \pi(\mathbf{s}, x^*), \; x^* \in \mathbf{A}. \tag{47}$$

A hotbooting technique as shown in Algorithm 1 exploits the experiences from large-scale UAV-aided VANET experiments to initialize the Q-value and the mixed-strategy π for each action-state pair to reduce the random exploration at the beginning of the repeated game and accelerate the learning process of the relay. More specifically, we consider ξ anti-jamming VANET transmission experiments in similar scenarios before the game. Each experiment that lasts K time slots, during which the UAV observes the current state, i.e., the anti-jamming transmission performance such as the BER of the OBU message in the VANET and then chooses the relay strategy with the mixed-strategy $\pi(\mathbf{s}, x)$ according to (44)-(46).

The initial Q-values $\mathbf{Q}^*$ and the mixed-strategy table π^* as the output of Algorithm 1 via ξ experiments are used to initialize the Q-values and mixed-strategy table in Algorithm 2, with $\mathbf{Q} = \mathbf{Q}^*$ and $\pi = \pi^*$. As shown in Algorithm 2, the UAV observes the current state $\mathbf{s}^{(k)}$ consisting of the channel quality of the OBU-RSU_1 link, the OBU-UAV link and the UAV-RSU_2 link and the BER of the OBU message at the previous time slot from the server. The relay decision $x^{(k)}$ is chosen according to (47). The utility of the UAV $u_U^{(k)}$ is evaluated and both the Q-function and mixed-strategy π are updated via (44)-(46) in each time slot. In this way, the UAV learns the jamming strategy in the anti-jamming transmission dynamic game and achieves the optimal relay strategy to improve the long-term anti-jamming transmission performance.



(a) BER of the OBU message

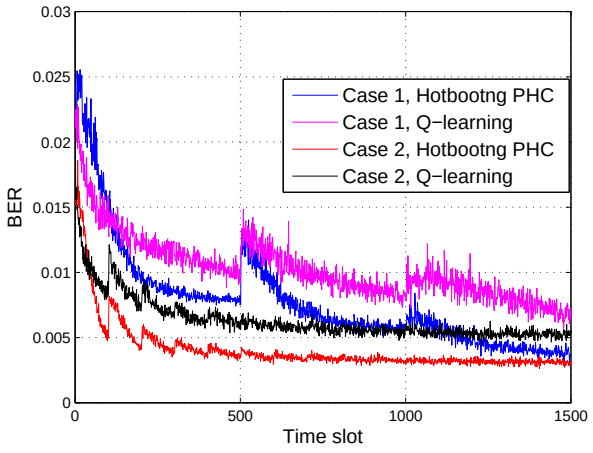


(b) Utility of the UAV

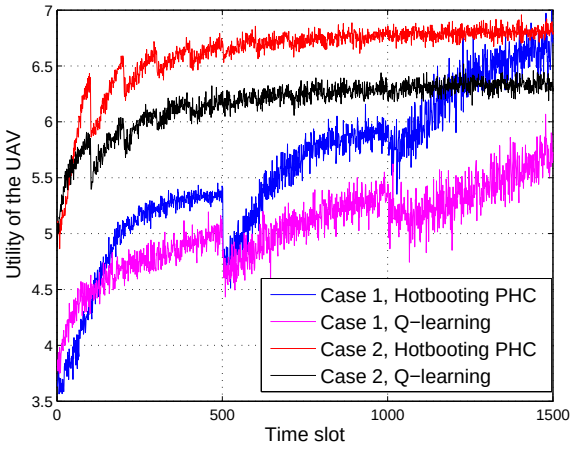
Fig. 5: The communication performance of the UAV-aided VANET in the dynamic game, with $C_U = 1$ and $C_J = 0.5$.

VII. Simulation Results

Simulations have been performed to evaluate the performance of the proposed UAV relay strategy in the dynamic game against smart jamming. In the simulations, both the UAV and the jammer were stationary and observed the ongoing VANET communication. The jammer applied greedy strategy to choose the jamming power regarding the expected reward u_J in (5) during the VANET communication in each time slot. The radio link between the jammer and RSU_2 was much worse than that between the jammer and RSU_1

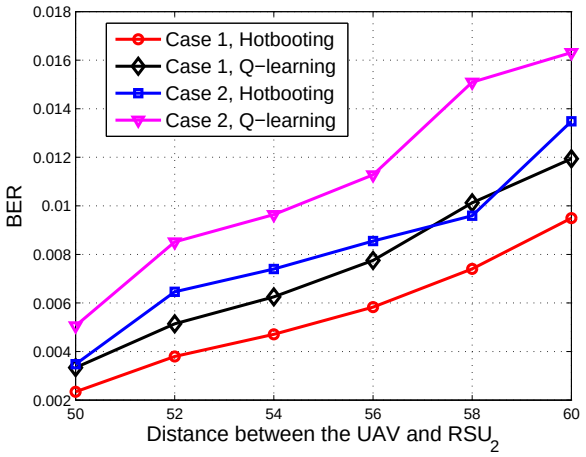


(a) BER of the OBU message

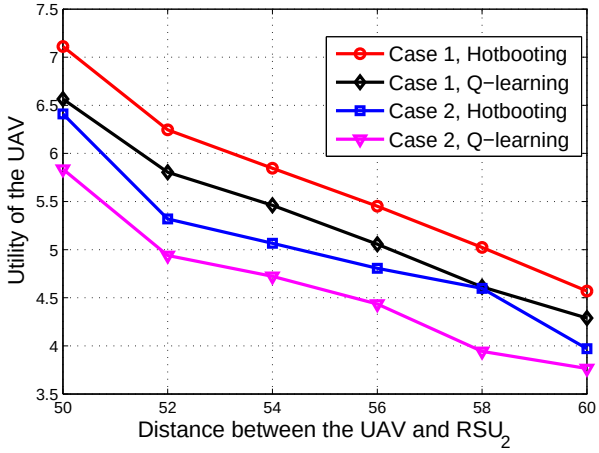


(b) Utility of the UAV

Fig. 6: Communication performance of the UAV-aided VANET in the dynamic game against the smart jammer that changes the jamming policy every 500 time slots in Case 1, with the channel conditions changing randomly every 100 time slots in Case 2.



(a) Average BER of the OBU messages



(b) Average utility of the UAV

Fig. 7: Communication performance of the UAV-aided VANET in the dynamic game with $d_4 = 60$ m in Case 1, and $d_4 = 50$ m in Case 2.

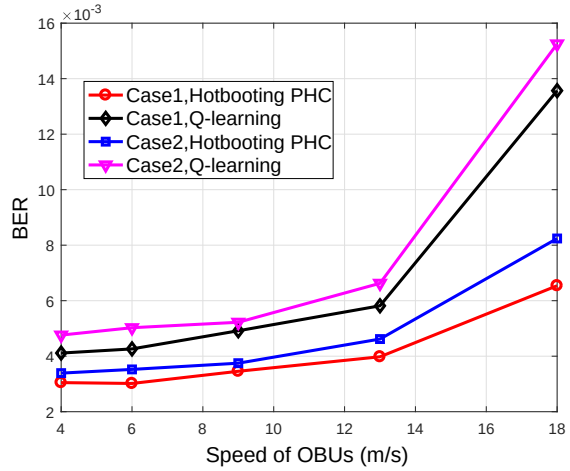
due to a longer distance, which leads to a lower channel gain. Unless specified otherwise, we set $P = 10$, $P_U = 2$, $C_U = 1$, $C_J = 0.5$, $\sigma = 0.1$, $h_3 = 0.4$ and $h_4 = 0.2$. More specifically, the transmit power of the UAV was 5 times higher than that of the OBU, with a unit relay energy cost. The jammer had a lower channel gain to the UAV than RSU_1 due to a longer distance, i.e., $h_3 = 0.4 > h_4 = 0.2$. Similar to the channel model in [33], the Doppler shift was consider in the OBU-RSU_1 and OBU-UAV channel models. The channel gain transition probability linearly increases with the moving speed of the OBU, and is given by

$$p_{i,m,n}^{(k)} = \begin{cases} \frac{\varphi v^{(k)}}{V}, & \text{if } (m,n) = (1,2) \text{ or } (N, N-1) \\ 1 - \frac{\varphi v^{(k)}}{V}, & \text{if } 1 \leq m = n \leq N \\ \frac{\varphi v^{(k)}}{2V}, & \text{if } 2 \leq m \leq N-1 \text{ and } n = m \pm 1 \\ 0, & \text{otherwise,} \end{cases} \tag{48}$$

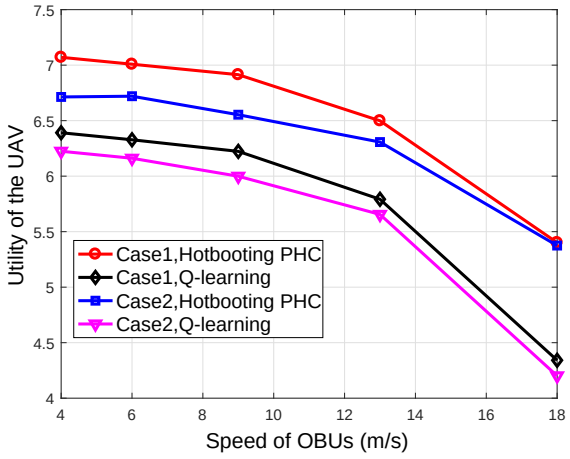
where φ denotes the impact of the environmental changes.

As shown in Fig. 5, the BER of the OBU message decreases with time, e.g., from 14.1‰ at the beginning of the game to 3‰ after 1500 time slots, which is about 65.7% lower than that of the Q-learning based strategy. This is because the PHC-based relay strategy increases the randomness of the UAV compared with Q-learning. The utility of the UAV increases by about 52.2% after 1500 time slots, which is about 42% higher than Q-learning. Our proposed scheme takes about 500 time slots (about 0.45 seconds) to converge to the optimal policy under static network environment and keeps using it unless the network state or the attack policy changes. The convergent speed of the hotbooting PHC-based relay is 56.3% faster than the standard PHC, due to the hotbooting technique formulates the emulated experiences.

The anti-jamming transmission performance of the dynamic game is shown in Fig. 6, in which a smart jammer changed its jamming power according to the expected learning results of the UAV relay every 500 time slots in Case 1. The channel conditions changed randomly every 100 time slots in Case



(a) Average BER of the OBU messages



(b) Average utility of the UAV

Fig. 8: Communication performance of the UAV-aided VANET in the dynamic game with $C_U = 0.5$ and $C_J = 0.1$ in Case 1, and $C_U = 1$ and $C_J = 0.5$ in Case 2.

2. The proposed strategy is more robust against the smart jammer, e.g., our proposed scheme reduces the BER of the OBU message in Case 1 by 40.9% and increases the utility of the UAV by 18.3% compared with Q-learning at time slot 1000. In Case 2, the BER of the OBU message of the hotbooting PHC-based relay is 37.5% lower than Q-learning at time slot 100. Our proposed scheme applies the hotbooting technique to accelerate the learning process in the jamming resistance, yielding a faster learning speed compared with the radio channel variations.

The BERs of the OBU messages averaged over 1500 time slots (approximately 1.36 seconds) are presented in Fig. 7, showing that the BER increases with the UAV-RSU$_2$ distance and decreases with the jammer-UAV distance. For example, the BER in Case 1 decreases from 9.4‰ to 2.3‰ as d_5 changes from 60 to 50 m, which is 30.4% lower than the Q-learning based strategy. The average utility of the UAV decreases with the UAV-RSU$_2$ distance and increases with the jammer-UAV distance, e.g., the average utility increases by 19% as d_4 changes from 50 to 60 m. That is because the UAV has a better radio link with OBU and RSU$_2$ due to the longer distance from the jammer.

The performance of the dynamic relay game is shown in Fig. 8, with $C_U = 0.5$ and $C_J = 0.1$ in Case 1, and $C_U = 1$ and $C_J = 0.5$ in Case 2. The average BER of the OBU messages increases with the OBU speed, e.g., from 3‰ to 6.54‰ as v changes from 4 to 18 m/s in Case 1, which is about 63% lower than Q-learning. That is because the high mobility of the OBU increases the transition probabilities of h_1 and h_2 and leads to an unstable channel condition. The average utility of the UAV decreases by about 24% as v changes from 4 to 18 m/s, which is about 21% lower than Q-learning. The average BER in Case 2 is higher than that in Case 1, because the UAV has less motivation to relay the OBU message with a higher transmission cost.

VIII. Conclusion

In this paper, we have formulated a UAV-aided VANET transmission game and derived the NEs of the game to disclose the impacts of the UAV transmit cost and the radio channel condition on communication performance of the VANET against smart jamming. A hotbooting PHC-based anti-jamming relay strategy has been proposed for a UAV to achieve the optimal relay policy without knowing the VANET model and the jamming model. Simulation results have shown that the proposed relay strategy can efficiently improve the anti-jamming transmission performance of the VANET. For example, this scheme reduces the BER of the OBU message in the VANET by 65.7% and increases the utility of the UAV by 42% compared with the Q-learning based relay strategy.

We will carry out a more in-depth game theoretic study of the anti-jamming UAV relay in VANETs in the future, which incorporates more jamming models, including a smart jammer that can choose both the jamming strength and the mobility. A backup anti-jamming transmission mechanism need to be provided to protect VANETs at the beginning stage of the learning process. In addition, we should develop new RL techniques with low computation and communication overhead to improve the anti-jamming communication performance of UAV-aided VANETs.

References

[1] H. Hartenstein and L. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 164–171, Jun. 2008.

[2] I. K. Azogu, M. T. Ferreira, J. A. Larcom, and H. Liu, "A new anti-jamming strategy for VANET metrics-directed security defense," in *Proc. IEEE Global Commun. Conf.*, pp. 1344–1349, Atlanta, GA, Dec. 2013.

[3] H. Sedjelmaci, S. M. Senouci, and N. Ansari, "Intrusion detection and ejection framework against lethal attacks in UAV-aided networks: A bayesian game-theoretic methodology," *IEEE Trans. Intell. Transportation Syst.*, pp. 1–11, Aug. 2016.

[4] Y. Zhou, N. Cheng, N. Lu, and X. Shen, "Multi-UAV-aided networks: Aerial-ground cooperative vehicular networking architecture," *IEEE Veh. Technol. Mag.*, vol. 10, no. 4, pp. 36–44, Dec. 2015.

[5] A. Sanjab, W. Saad, and T. Başar, "Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game," *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017.

[6] L. Xiao, *Anti-jamming transmissions in cognitive radio networks*. Springer, 2015.

[7] P. Zhan, K. Yu, and A. L. Swindlehurst, "Wireless relay communications with unmanned aerial vehicles: Performance and optimization," *IEEE Trans. Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 2068–2085, Jul. 2011.

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

[9] G. Han, L. Xiao, and H. V. Poor, "Two-dimensional anti-jamming communication based on deep reinforcement learning," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, New Orleans, LA, Mar. 2017.

[10] M. Bowling and M. Veloso, "Rational and convergent learning in stochastic games," in *Proc. Int. Joint Conf. Artificial Intelligence*, pp. 1021–1026, Seattle, WA, Aug. 2001.

[11] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345 – 1359, Oct. 2010.

[12] J. Peng and L. Cheng, "A distributed mac scheme for emergency message dissemination in vehicular ad hoc networks," *IEEE Trans. Vehicular Technology*, vol. 56, no. 6, pp. 3300–3308, Nov. 2007.

[13] K. A. Hafeez, L. Zhao, J. W. Mark, X. Shen, and Z. Niu, "Distributed multichannel and mobility aware cluster-based mac protocol for vehicular ad-hoc networks (vanet)," *IEEE Tran. on Vehicular Technology*, vol. 62, no. 8, pp. 3886–3902, Oct. 2013.

[14] J. Zhao and G. Cao, "Vadd: Vehicle-assisted data delivery in vehicular ad hoc networks," *IEEE Trans. Vehicular Technology*, vol. 57, no. 3, pp. 1910–1922, May 2008.

[15] A. Hamieh, J. Ben-Othman, and L. Mokdad, "Detection of radio interference attacks in VANET," in *Proc. IEEE Global Commun. Conf.*, pp. 1–5, Honolulu, Hawaii, Nov. 2009.

[16] T. Alpcan and S. Buchegger, "Security games for vehicular networks," *IEEE Trans. Mobile Computing*, vol. 10, no. 2, pp. 280–290, Feb. 2011.

[17] C. Dixon and E. W. Frew, "Optimizing cascaded chains of unmanned aircraft acting as communication relays," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 5, pp. 883–898, Jun. 2012.

[18] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 4983 – 4996, Dec. 2016.

[19] Q. Wang, Z. Chen, W. Mei, and J. Fang, "Improving physical layer security using UAV-enabled mobile relaying," *IEEE Wireless Comm. Letters*, Mar. 2017.

[20] J. Kosmerl and A. Vilhar, "Base stations placement optimization in wireless networks for emergency communications," in *Proc. IEEE Int. Conf. Commun.*, pp. 200–205, Sydney, Australia, Jun. 2012.

[21] G. Tuna, T. V. Mumcu, K. Gulez, V. C. Gungor and H. Erturk, "Unmanned aerial vehicle-aided wireless sensor network deployment system for post-disaster monitoring," in *Emerging Intelligent Computing Technology and Applications*, vol. 304, pp. 298–305, Jul. 2012.

[22] T. A. Johansen, A. Zolich, and T. Hansen, "Unmanned aerial vehicle as communication relay for autonomous underwater vehicle-Field tests," in *Proc. IEEE Global Commun. Conf.*, pp. 1469–1474, Austin, TX, Dec. 2014.

[23] J. Ueyama, H. Freitas, B. S. Faiçal, et al., "Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 81–87, Dec. 2014.

[24] M. Dong, K. Ota, M. Lin, and et al., "UAV-assisted data gathering in wireless sensor networks," *J. Supercomput.*, vol. 70, no. 3, pp. 1142–1155, Dec. 2014.

[25] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, Jun. 2016.

[26] O. Puñal, A. Aguiar, and J. Gross, "In VANETs we trust?: Characterizing RF jamming in vehicular networks," in *Proc. ACM Int. Workshop Veh. Inter-Netw. Syst. Appl.*, pp. 83–92, Ambleside, UK, Jun. 2012.

[27] Z. Lu, W. Wang, and C. Wang, "Modeling, evaluation and detection of jamming attacks in time-critical wireless applications," *IEEE Trans. Mobile Computing*, vol. 13, no. 8, pp. 1746–1759, Aug. 2014.

[28] A. Benslimane and H. Nguyen-Minh, "Jamming attack model and detection method for beacons under multichannel operation in vehicular networks," *IEEE Trans. Vehicular Technology*, Dec. 2016.

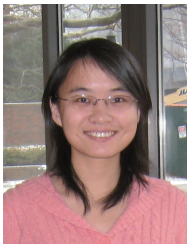
[29] C. Long, Q. Zhang, B. Li, H. Yang, and X. Guan, "Non-cooperative power control for wireless ad hoc networks with repeated games," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 6, pp. 1101–1112, Aug. 2007.

[30] X. Lu, D. Xu, L. Xiao, L. Wang, and W. Zhuang, "Anti-jamming communication game for UAV-aided VANETs," *submitted to IEEE Global Commun. Conf.*, Singapore, Dec. 2017.

[31] V. Erceg, L. J. Greenstein, S. Y. Tjandra, and et al., "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1205 – 1211, Jul. 1999.

[32] R. Palat, A. Annamalau, and J. Reed, "Cooperative relaying for ad-hoc ground networks using swarm UAVs," in *Proc. IEEE Military Commun. Conf.*, pp. 1588–1594, Atlantic, NJ, Oct. 2005.

[33] L. Xiao, T. Chen, C. Xie, H. Dai, and H. V. Poor, "Mobile crowdsensing games in vehicular networks," *IEEE Trans. Vehicular Technology*, Jan. 2017.



Liang Xiao (M'09, SM'13) is currently a Professor in the Department of Communication Engineering, Xiamen University, Fujian, China. She has served as an associate editor of IEEE Trans. Information Forensics and Security and guest editor of IEEE Journal of Selected Topics in Signal Processing. She is the recipient of the best paper award for 2016 INFOCOM Big Security WS and 2017 ICC. She received the B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications, China, in 2000, the M.S. degree in electrical engineering from Tsinghua University, China, in 2003, and the Ph.D. degree in electrical engineering from Rutgers University, NJ, in 2009. She was a visiting professor with Princeton University, Virginia Tech, and University of Maryland, College Park.

Xiaozhen Lu received the B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2017. She is currently pursuing the PhD. degree with the Department of Communication Engineering, Xiamen University, Xiamen, China. Her research interests include network security and wireless communications.

Dongjin Xu received the B.S. degree in communication engineering from Xiamen University, Xiamen, China, in 2016, where she is currently pursuing the M.S. degree with the Department of Communication Engineering. Her research interests include network security and wireless communications.

Lei Wang received the M.Sc. degree and the Ph.D. degree in Telecommunications and Information Engineering from Nanjing University of Posts and Telecommunications, China, in 2007 and 2010, respectively. From 2012 to 2013, he was a Postdoctoral research fellow at the Department of Electrical Engineering, Columbia University, USA. He is currently an associate professor at the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China. His research interests include millimeter wave wireless communications, device to device communications, physical-layer security, signal processing for communications, cognitive wireless networks, and random matrix theory.

Weihua Zhuang (M'93-SM'01-F'08) has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, since 1993, where she is a Professor and a Tier I Canada Research Chair in Wireless Communication Networks. She is the recipient of 2017 Technical Recognition Award from IEEE Communications Society Ad Hoc & Sensor Networks Technical Committee, one of 2017 ten N2Women (Stars in Computer Networking and Communications), and a co-recipient of several best paper awards from IEEE conferences. Dr. Zhuang was the Editor-in-Chief of IEEE Transactions on Vehicular Technology (2007-2013), Technical Program Chair/Co-Chair of IEEE VTC Fall 2017 and Fall 2016, and the Technical Program Symposia Chair of the IEEE Globecom 2011. She is a Fellow of the IEEE, the Royal Society of Canada, the Canadian Academy of Engineering, and the Engineering Institute of Canada. Dr. Zhuang is an elected member in the Board of Governors and VP Publications of the IEEE Vehicular Technology Society. She was an IEEE Communications Society Distinguished Lecturer (2008-2011).