Anti-Jamming Communication Game for UAV-aided VANETs

Xiaozhen Lu*, Dongjin Xu*, Liang Xiao*, Lei Wang†, Weihua Zhuang‡
*Dept. of Communication Engineering, Xiamen University, Xiamen, China. Email: lxiao@xmu.edu.cn
†Dept. of TIE, Nanjing University of Posts and Telecommunications, Nanjing, China. Email: wanglei@njupt.edu.cn
‡Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. Email: wzhuang@uwaterloo.ca

***Abstract*—Vehicular ad-hoc networks (VANETs) are vulnerable to jamming attacks, and frequency hopping-based anti-jamming techniques are not always applicable in VANETs due to the high mobility of the onboard units (OBUs) especially under a large scale network topology. In this paper, we use unmanned aerial vehicles (UAVs) to deal with VANET jamming, especially smart jamming that changes the jamming policy based on the ongoing communication status of the VANET. More specifically, the UAV relays the data of OBUs to another roadside unit (RSU) with a better transmission condition if the serving RSU is located in a heavily jammed area. The interactions between the UAV and the jammer are formulated as an anti-jamming UAV relay game, in which the UAV decides whether or not to relay the data of the OBU to another RSU that is far away from the jammer, and the latter chooses the jamming power. The Nash equilibria (NE) of the game are derived to reveal how the best UAV relay strategy depends on the transmission cost and the radio channel model. A hotbooting policy hill climbing (PHC)-based UAV relay strategy is proposed to address jamming in the dynamic UAV-aided VANET game without the knowledge of network model and jamming model. Simulation results show that the proposed relay strategy can efficiently reduce the bit error rate (BER) of OBU data and thus increase the utility of VANET in comparison with a Q-learning based scheme.**

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

I. Introduction

Vehicular ad-hoc networks (VANETs) supporting vehicle-to-vehicle (V2V) communications and/or vehicle-to-infrastructure (V2I) communications can improve road safety, vehicular transport efficiency, and booming information/entertainment applications of onboard units (OBUs) on vehicles [1]. The high mobility of OBUs and the large-scale network with infrastructures such as roadside units (RSUs) make VANET vulnerable to jamming [2], especially from smart jammers who send faked or replayed signals with the goal to block the ongoing transmissions between the OBUs and the serving RSUs with flexible jamming strategies and strong radio sensing capabilities.

The jamming resistance of VANETs can be significantly improved by using unmanned aerial vehicles (UAVs), which are generally faster to deploy and have better radio communication channels with both OBUs and RSUs, as compared with the RSUs located at a known location due to the line-of-sight (LOS) links and a smaller path-loss exponent [3]. Therefore, we consider a UAV-aided VANET to address jamming attacks, in which the UAVs help relay the data in the VANET to improve the signal to interference plus noise ratio (SINR) of the signals sent from or to the OBUs, and thus reduce the bit error rate (BER) of the data, especially if the serving RSU is deeply jammed.

Game theory has been used to study jamming attacks in wireless networks [4]–[9]. However, to our best knowledge, there are limited studies on jamming in VANETs, especially the UAV relay in VANETs. In this paper, we formulate an anti-jamming UAV relay game to study the jamming resistance in a UAV-aided VANET, in which the UAV decides whether or not to relay the data of an OBU to an RSU, and the smart jammer chooses its jamming power. This game assumes an air-to-ground channel model that consists of path loss, log-normal shadowing and Rayleigh fading [10]. The Nash equilibria (NEs) of the game are derived and the conditions that each NE exists are provided to disclose the impacts of the transmission cost and the radio channel condition on the jamming resistance of VANETs.

As a smart jammer determines its jamming power based on the previous VANET security performance, the UAV relay decisions in the dynamic game can be formulated as a Markov decision process (MDP). Reinforcement learning such as Q-learning can provide an optimal strategy via trials in MDPs [11], [12]. Therefore, we propose a reinforcement learning based UAV relay strategy to resist jamming without the knowledge of jamming model in the dynamic game. More specifically, we design a UAV relay scheme based on policy hill climbing (PHC), which increases the randomness compared with Q-learning at the beginning of exploration, and thus fools the opponent.

The random exploration of standard PHC based on all-zero Q-value initialization is quite inefficient in some environments, requiring exponentially more data than the optimal policy. By exploiting the experience in similar scenarios, we design a hotbooting technique to initialize the values of the Q-function and the probability π for each action-state pair to speed up the learning process. Simulation results show that

This work was supported in part by National Natural Science Foundation of China under Grant 61671396, CCF-Venustech Hongyan Research Initiative (2016-010) and the Major Projects of the Natural Science Foundation of the Jiangsu Higher Education Institutions (16KJA510004)

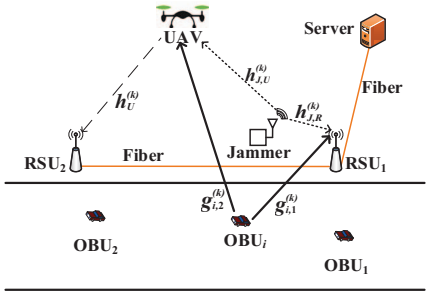


Fig. 1: Illustration of a UAV-aided VANET against jamming attacks.

the proposed relay strategy can efficiently reduce the BER of the data in the VANET with quadrature phase-shift keying (QPSK) and thus increase the utility of VANET.

The contributions of this work can be summarized as:

- We formulate an anti-jamming relay game to study the interactions between a UAV and a smart jammer in the UAV-aided VANET and provide conditions under which the NEs of the game exist.
- We propose a hotbooting PHC-based UAV relay strategy to resist smart jamming in the VANET without the knowledge of channel model and jamming model, and perform simulations to verify its performance gain over the Q-learning based relay strategy with a lower BER of the data and a higher utility.

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

II. Related Work

The jamming detection model formulated in [13] provides the correlation between jamming detection error and data reception time of the VANET. The jamming defense game as formulated in [14] deals with the allocation of the VANET defense resources. The ambient noise immunity based anti-jamming algorithm developed in [15] adjusts the false detection threshold to improve the packets delivery rate in VANETs. The hideaway strategy as proposed in [2] determines when to keep silent based on the packet send ratio to improve jamming resistance.

The UAV-aided wireless sensor network (WSN) as investigated in [16] can reduce the packet loss and power consumption of the WSN with node failures. The field tests as conducted in [17] investigates the impact of the UAV altitude and radius on the communication quality in autonomous underwater vehicle. The optimal placement of UAVs as analyzed in [18] enhances the coverage of public safety communications. The tradeoff between the coverage and the time required to cover the entire target area of the UAV has been analyzed in [19] to determine the number of stop points of the UAV in an underlaid device-to-device communication network. The UAV-enabled mobile relaying system as investigated in [20] uses the difference-of-concave program to optimize the transmit power of the mobile device and the relay to maximize the secrecy rate.

III. System Model

A. Network Model

In this work, we consider a VANET consisting of several OBUs, several RSUs, a UAV as relay, and the system server, as illustrated in Fig. 1. RSUs at the fixed locations are connected with each other and the server via fibers. Equipped with sensors, such as cameras and global positioning system receivers, OBUs can gather the sensing information and transmit it to the server via RSUs. In a time-slotted system, each OBU is assumed to send a signal to the serving RSU, denoted by RSU$_1$, over the radio channel with the channel power gain at time slot k denoted by $g_{i,1}^{(k)}$. The transmission is vulnerable to the jamming signals sent by a smart jammer. The jammer is assumed to be too far away from another RSU denoted by RSU$_2$ to block the transmission.

In the system, both the UAV and RSU$_1$ connect with one OBU in different time slots. The UAV connects to the server via RSU$_2$. The jammer chooses its jamming power $y \in [0, P_J^M]$, where P_J^M is the maximum power. According to the channel quality and the BER of the data sent by OBU$_i$, the UAV decides whether or not to relay the data of OBU$_i$ to RSU$_2$, which is denoted by $x_i \in \{0, 1\}$. The UAV relays the data if $x_i = 1$, and keeps silent otherwise. The system server can gather the sensing data sent from OBU$_i$ via RSU$_1$ or the UAV.

The SINR of the signals received by r ($r = R$ for RSU$_1$ or $r = U$ for the UAV) that sent from OBU$_i$, and the SINR of the signals received by RSU$_2$ that sent from the UAV at time slot k, are denoted by $\rho_{i,r}^{(k)}$ and $\rho_{U,R}^{(k)}$, respectively.

Let $p_e(m)$ denote the BER of the signal, where m is the signal-to-noise ratio per bit, or E_b/N_0. The BER $P_e^{(k)}$ of the signal sent by OBU$_i$ at time slot k based on $\rho_{i,r}^{(k)}$, and $\rho_{U,R}^{(k)}$, is given by

$$P_e^{(k)} = \min\left(p_e\left(\rho_{i,R}^{(k)}\right), p_e\left(\min\left(\rho_{i,U}^{(k)}, \rho_{U,R}^{(k)}\right)\right)\right). \quad (1)$$

The transmission cost of the UAV at time slot k is denoted by $C_U^{(k)}$, and that of the jammer is denoted by $C_J^{(k)}$. For simplicity, we assume a constant noise power denoted by σ and a fixed transmit power $P_i^{(k)}$ (or $P_U^{(k)}$) at OBU$_i$ (or the UAV) .

B. Channel Model

Due to the mobility of OBU$_i$, the radio channel state of the link from OBU$_i$ to its serving RSU or the UAV changes over time. For simplicity, the channel gain of OBU$_i$ to the UAV at time slot k is denoted by $g_{i,2}^{(k)}$. Let $\mathbf{g}_i^{(k)} = \left[g_{i,j}^{(k)}\right]_{1 \le j \le M} \in$

$\{G_a\}_{0\le a\le N_j}^{M}$, with each element quantized into N_j+1 levels. The channel power gain of each OBU is modeled as a Markov chain with N_j+1 states, in which the transfer probability of $g_{i,j}$ from G_m to G_n at time slot k is denoted by $p_{m,n,i}^{(k)}$, and is given by

$$p_{m,n,i}^{(k)} = \Pr\left(g_{i,j}^{(k)} = G_n | g_{i,j}^{(k-1)} = G_m\right). \tag{2}$$

Due to the location distribution of the OBU$_i$-RSU$_1$ and OBU$_i$-UAV radio links, $g_{i,1}^{(k)}$ and $g_{i,2}^{(k)}$ are assumed to be independent, and changes every time slot. The radio link between the UAV and OBU$_i$ has a higher path loss compared with that between RSU$_1$ and OBU$_i$ due to the longer distance. The UAV transmission can be congested if the distance from the UAV to OBU$_i$ is too long, while the UAV coverage is large. The UAV is deployed at a proper distance, which provides full coverage for the target area and transmission.

Let $\mathbf{h}_J^{(k)} = \left[h_{J,R}^{(k)}, h_{J,U}^{(k)}\right]$ denote the channel power gain vector of the jammer at time slot k, where $h_{J,R}^{(k)}$ (or $h_{J,U}^{(k)}$) is the channel power gain between the jammer and RSU$_1$ (or the UAV). The channel power gain between the UAV and RSU$_2$ at time slot k is denoted by $h_U^{(k)}$.

IV. ANTI-JAMMING TRANSMISSION GAME

A. Anti-Jamming UAV Relay 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 data of OBU$_i$ to RSU$_2$, $x_i \in \{0,1\}$, 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 transmission cost, which is given by

$$u_U^{(k)}(x_i, y) = x_i \max\left(\frac{P_i^{(k)} g_{i,1}^{(k)}}{\sigma + y h_{J,R}^{(k)}}, \min\left(\frac{P_i^{(k)} g_{i,2}^{(k)}}{\sigma + y h_{J,U}^{(k)}}, \frac{P_U^{(k)} h_U^{(k)}}{\sigma}\right)\right) - x_i C_U^{(k)} + \frac{(1-x_i) P_i^{(k)} g_{i,1}^{(k)}}{\sigma + y h_{J,R}^{(k)}}. \tag{3}$$

In (3), the first term in the max function represents the SINR of the signal sent by OBU$_i$ and received by RSU$_1$ at time slot k, and the second term represents the SINR of weaker signal over the OBU$_i$-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

$$u_J^{(k)}(x_i, y) = -x_i \max\left(\frac{P_i^{(k)} g_{i,1}^{(k)}}{\sigma + y h_{J,R}^{(k)}}, \min\left(\frac{P_i^{(k)} g_{i,2}^{(k)}}{\sigma + y h_{J,U}^{(k)}}, \frac{P_U^{(k)} h_U^{(k)}}{\sigma}\right)\right) + x_i C_U^{(k)} + \frac{(x_i-1) P_i^{(k)} g_{i,1}^{(k)}}{\sigma + y h_{J,R}^{(k)}} - y C_J^{(k)}. \tag{4}$$

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

B. NE of the Game

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

$$u_U(x_i^*, y^*) \ge u_U(x_i, y^*),\ \forall x_i \in \{0,1\} \tag{5}$$

$$u_J(x_i^*, y^*) \ge u_J(x_i^*, y),\ \forall y \in \left[0, P_J^M\right]. \tag{6}$$

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

$$C_J \sigma^2 > P_i g_{i,1} h_{J,R} \tag{7}$$

and

$$C_U \sigma \ge \min\left(P_U h_U - P_i g_{i,1}, P_i g_{i,2} - P_i g_{i,1}\right) \tag{8}$$

or

$$P_i g_{i,1} > \max(P_U h_U, P_i g_{i,2}). \tag{9}$$

Proof: By (3), if (8) holds, we have

$$u_U(1,0) = \frac{P_U h_U}{\sigma} - C_U \le \frac{P_i g_{i,1}}{\sigma} = u_U(0,0). \tag{10}$$

Thus, (5) holds for $(x_i^*, y^*) = (0,0)$.

By (4), we have

$$u_J(0,y) = -\frac{P_i g_{i,1}}{\sigma + y h_{J,R}} - y C_J \tag{11}$$

$$\partial u_J(0,\tilde{y})/\partial y = \frac{P_i g_{i,1} h_{J,R}}{(\sigma + \tilde{y} h_{J,R})^2} - C_J = 0 \tag{12}$$

and $\partial^2 u_J/\partial y^2 \le 0$, indicating that if (7) holds, $u_J(0,y)$ is concave with respect to (w.r.t.) y, and $\tilde{y} = \sqrt{P_i g_{i,1}/C_J h_{J,R}} - \sigma/h_{J,R} < 0$. Thus,

$$u_J(0,0) = -\frac{P_i g_{i,1}}{\sigma} \ge -\frac{P_i g_{i,1}}{\sigma + y h_{J,R}} - y C_J = u_J(0,y), \forall y. \tag{13}$$

Thus, (6) holds for $(x_i^*, y^*) = (0,0)$, indicating that (0, 0) is an NE of the game. Similarly, we can prove that (0, 0) is an NE of the game if (7) and (9) hold. ■

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 OBU$_i$ data in the the VANET with QPSK. By (1), we have

$$P_e = \frac{1}{2}\min\left(\operatorname{erfc}\left(\sqrt{\frac{\rho_{i,R}}{2}}\right), \operatorname{erfc}\left(\sqrt{\frac{\min(\rho_{i,U}, \rho_{U,R})}{2}}\right)\right). \tag{14}$$

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

by

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\min(P_i g_{i,2}, P_U h_U)}{2\sigma}}\right). \quad (15)$$

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

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

Remark: Both the jammer and the UAV keep silent to reduce the power consumption if the jamming cost and the transmission cost are high as shown in (7) and (8) or the link between OBU$_i$ and RSU$_1$ is in a good condition as shown in (9). In the case, the BER of OBU$_i$ signal decreases with the transmit power of OBU$_i$ as shown in (15) and (16).

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

$$C_U\sigma + P_i g_{i,1} \leq P_U h_U < \frac{P_i g_{i,2}\sigma}{\sigma + P_J^M h_{J,U}} \quad (17)$$

or

$$\begin{cases} \dfrac{C_U\sigma}{g_{i,2} - g_{i,1}} \leq P_i < \min\left(\dfrac{C_J\sigma^2}{g_{i,1}h_{J,U}}, \dfrac{P_U h_U}{g_{i,2}}\right) & (18a) \\ h_{J,U} < h_{J,R}. & (18b) \end{cases}$$

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

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{P_U h_U}{2\sigma}}\right). \quad (19)$$

Corollary 4. *If (18) holds, the BER of OBU$_i$ signal in the UAV-aided VANET with QPSK at the NE of the game is given by*

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{P_i g_{i,2}}{2\sigma}}\right). \quad (20)$$

Remark: The UAV relays for OBU$_i$ to improve its performance if the transmission cost is low and the link between OBU$_i$ 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 (18a).

Theorem 3. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $\left(0, \sqrt{P_i g_{i,1}/C_J h_{J,R}} - \sigma/h_{J,R}\right)$*, if*

$$\frac{P_i g_{i,1} h_{J,R}}{(\sigma + P_J^M h_{J,R})^2} \leq C_J \leq \min\left(\frac{P_i g_{i,1} h_{J,R}}{\sigma^2}, \max\left(\frac{P_U^2 h_U^2 h_{J,R}}{P_i g_{i,1}\sigma^2}, \frac{P_i h_{J,R}(g_{i,2}h_{J,R} - g_{i,1}h_{J,U})^2}{g_{i,1}\sigma^2(h_{J,R} - h_{J,U})^2}\right)\right)$$

and

$$C_U \geq \min\left(\frac{P_U h_U}{\sigma} - \frac{P_i g_{i,1}}{\sigma + P_J^M h_{J,R}}, \frac{P_i(g_{i,2} - g_{i,1})}{\sigma}\right) \quad (21)$$

or

$$P_i g_{i,1} C_J \sigma^2 \geq \max\left(P_U^2 h_U^2 h_{J,R}, \frac{P_i^2 g_{i,2}^2 h_{J,R}^3 \sigma^2 C_J}{\left(\sigma(h_{J,R} - h_{J,U})\sqrt{C_J} + h_{J,U}\sqrt{P_i g_{i,1} h_{J,R}}\right)^2}\right). \quad (22)$$

Theorem 4. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $\left(1, \sqrt{P_i g_{i,2}/C_J h_{J,U}} - \sigma/h_{J,U}\right)$*, if*

$$\begin{cases} C_U \leq \dfrac{P_i g_{i,2}}{\sigma + P_J^M h_{J,U}} - \dfrac{P_i g_{i,1}}{\sigma + P_J^M h_{J,R}} & (23a) \\ \max\left(\dfrac{P_i g_{i,1} h_{J,U}}{(\sigma + P_J^M h_{J,U})^2}, \dfrac{P_i h_{J,U}(g_{i,1}h_{J,U} - g_{i,2}h_{J,R})^2}{g_{i,2}\sigma^2(h_{J,U} - h_{J,R})^2}\right) \\ \quad \leq C_J \leq \min\left(\dfrac{P_i g_{i,1} h_{J,U}}{\sigma^2}, \dfrac{P_U^2 h_U^2 h_{J,U}}{P_i g_{i,2}\sigma^2}\right). & (23b) \end{cases}$$

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

$$C_J(\sigma + P_J^M h_{J,R})^2 < P_i g_{i,1} h_{J,R} \quad (24)$$

and

$$C_U \geq \min\left(\frac{P_U h_U}{\sigma} - \frac{P_i g_{i,1}}{\sigma + P_J^M h_{J,R}}, \frac{P_i g_{i,2}}{\sigma + P_J^M h_{J,U}} - \frac{P_i g_{i,1}}{\sigma + P_J^M h_{J,R}}\right) \quad (25)$$

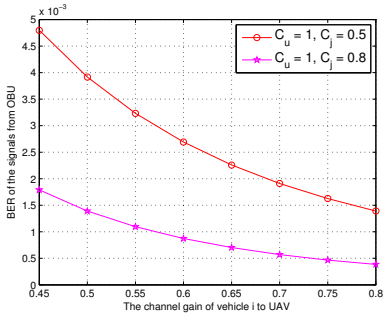
or

$$P_i g_{i,1}\sigma > \max\left(P_U h_U(\sigma + P_J^M h_{J,R}), P_i g_{i,2}\sigma\right). \quad (26)$$

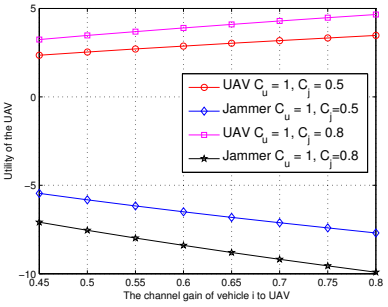
Theorem 6. *The anti-jamming transmission game in the UAV-aided VANET has an NE* $(1, P_J^M)$*, if*

$$\begin{cases} C_U \leq \dfrac{P_i g_{i,2}}{\sigma + P_J^M h_{J,U}} - \dfrac{P_i g_{i,1}}{\sigma + P_J^M h_{J,R}} & (27a) \\ \dfrac{C_J(\sigma + P_J^M h_{J,U})^2}{g_{i,1} h_{J,U}} < P_i \leq \dfrac{P_U h_U}{g_{i,2}} & (27b) \\ g_{i,2}\sigma > g_{i,1}(\sigma + P_J^M h_{J,U}). & (27c) \end{cases}$$

The performance at the NE of the game is evaluated with $P_i = 10$, $P_U = 2$, $\sigma = 0.1$, $g_{i,1} = 0.2$, $h_{J,R} = 0.4$, $h_{J,U} = 0.2$ and $h_U = 0.3$. As shown in Fig. 2, the BER decreases with the channel gain $g_{i,2}$ and the jamming cost, e.g., it decreases by 62.5% as C_J changes from 0.5 to 0.8. Consequently, the utility of the UAV increases with the channel gain $g_{i,2}$, while the utility of the jammer decreases with it. For example, the utility of the UAV increases by 47.5% as the channel gain $g_{i,2}$ changes from 0.45 to 0.8. In addition, the utility of the UAV increases with the jamming cost, e.g., it increases by 38% as C_J changes from 0.5 to 0.8, because the jammer is less motivated with a higher jamming cost.



(a) BER of the signals from OBU



(b) Utility of the UAV

Fig. 2: Performance of the game vs. the channel gain of OBU$_i$ to UAV, $g_{i,2}$, with $P_i = 10$, $P_U = 2$, $\sigma = 0.1$, $g_{i,1} = 0.2$, $h_{J,R} = 0.4$, $h_{J,U} = 0.2$ and $h_U = 0.3$.

V. 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. The UAV relay decision has impacts on the future jamming policy of the smart jammer, and thus can be formulated as an MDP. Therefore, the UAV can apply reinforcement learning techniques such as PHC to derive its optimal strategy via trials without the knowledge of jamming model.

We design a hotbooting technique that exploits the training data obtained in advance from large-scale experiments to initialize the Q-value. The hotbooting PHC-based relay scheme can decrease the useless random explorations and accelerate the learning speed in the relay for VANET compared with PHC. More specifically, the UAV performs N experiments in similar scenarios at the start of the game. In each experiment, the relay strategy is randomly selected at a given state to observe the resulting anti-jamming transmission performance, such as the BER of OBUs signals in the VANET.

In the dynamic game, the UAV decides whether or not to relay the OBU$_i$ data based on the system state at time slot k denoted by $\mathbf{s}^{(k)}$ that consists of the link quality between the UAV and OBU$_i$, that between RSU$_1$ and OBU$_i$, the SINR between RSU$_2$ and the UAV, and the BER of the OBU$_i$ data at the previous time slot, i.e., $\mathbf{s}^{(k)} = \left[\rho_{i,R}^{(k-1)}, \rho_{i,U}^{(k-1)}, \rho_{U,R}^{(k-1)}, P_e^{(k-1)}\right]$.

Let $\alpha \in (0,1]$ denote the learning rate of the UAV and $\delta \in [0,1]$ denote the discount factor of the UAV. The Q-function

Algorithm 1. Hotbooting PHC-based UAV relay strategy.

Initialize α, β, δ, $\mathbf{s}^{(0)}$ and $\mathbf{A}$.
$\mathbf{Q} = \mathbf{Q}^*$, $\mathbf{V} = \mathbf{0}$, $\pi = \pi^*$.
for $k = 1, 2, 3, ...$
  Choose $x_i^{(k)}$ with the probability $\pi\left(\mathbf{s}^{(k)}, x_i^{(k)}\right)$;
  if $x_i^{(k)} = 1$
    Relay the data of OBU$_i$ to RSU$_2$;
  End if
  Collect the SINR $\rho_{i,R}, \rho_{i,U}, \rho_{U,R}$, and the BER P_e from server;
  Obtain utility $u_U^{(k)}$;
  Update $Q\left(\mathbf{s}^{(k)}, x_i^{(k)}\right)$ via (28);
  Update $V\left(\mathbf{s}^{(k)}\right)$ via (29);
  Update $\pi\left(\mathbf{s}^{(k)}, x_i^{(k)}\right)$ via (30);
  $\mathbf{s}^{(k+1)} = \left[\rho_{i,R}^{(k)}, \rho_{i,U}^{(k)}, \rho_{U,R}^{(k)}, P_e^{(k)}\right]$.
End for

of the UAV with its action $x_i^{(k)}$ at state $\mathbf{s}^{(k)}$ is denoted by $Q\left(\mathbf{s}^{(k)}, x_i^{(k)}\right)$, and is updated according to iterative Bellman equation as follows:

$$Q\left(\mathbf{s}^{(k)}, x_i^{(k)}\right) \leftarrow (1-\alpha) Q\left(\mathbf{s}^{(k)}, x_i^{(k)}\right) + \alpha\left(u_U\left(\mathbf{s}^{(k)}, x_i^{(k)}\right) + \delta V\left(\mathbf{s}^{(k+1)}\right)\right) \quad (28)$$

$$V\left(\mathbf{s}^{(k)}\right) = \max_{x \in \{0,1\}} Q\left(\mathbf{s}^{(k)}, x\right) \quad (29)$$

where $V\left(\mathbf{s}^{(k)}\right)$ maximizes the Q value over the relay strategy.

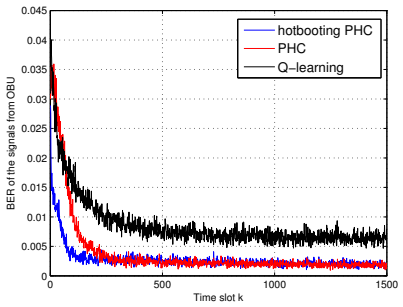
The UAV updates the mixed strategies table π by increasing the probability corresponding to the highest valued action by $\beta \in (0,1]$, and decreasing other probabilities by $-\beta/(|\mathbf{A}|-1)$, i.e.,

$$\pi\left(\mathbf{s}^{(k)}, x_i^{(k)}\right) \leftarrow \pi\left(\mathbf{s}^{(k)}, x_i^{(k)}\right) + \begin{cases} \beta, & x_i = \max\limits_{x \in \{0,1\}} Q\left(\mathbf{s}^{(k)}, x\right) \\ \frac{-\beta}{|\mathbf{A}|-1}, & \text{o.w.} \end{cases} \quad (30)$$

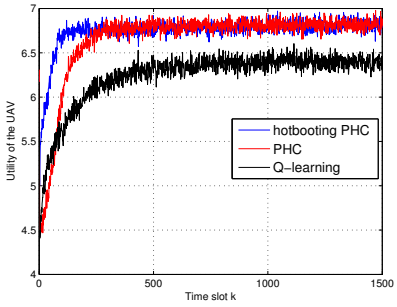
Before the game, we set $\mathbf{Q} = \mathbf{Q}^*$ and $\pi = \pi^*$ in the hotbooting PHC-based anti-jamming UAV relay algorithm, as summarized in Algorithm 1.

VI. Simulation Results

Simulations have been performed to evaluate the performance of the proposed UAV relay strategy in the dynamic game against smart jamming with $P_i = 10$, $P_U = 2$, $\sigma = 0.1$, $h_{J,R} = 0.4$, $h_{J,U} = 0.2$, $C_U = 1$ and $C_J = 0.5$. In the simulations, the UAV was stationary and the smart jammer chose its policy to maximize its immediate utility estimated according to the jamming history. The path loss of the radio link between the jammer and the UAV was much worse than that between the jammer and RSU$_1$ due to a shorter distance, which leads to a lower channel gain.



(a) BER of the signals from OBU



(b) Utility of the UAV

Fig. 3: Performance of the learning-based anti-jamming UAV relay schemes in the dynamic anti-jamming VANET communication game, with $P_i = 10$, $P_U = 2$, $\sigma = 0.1$, $h_{J,R} = 0.4$, $h_{J,U} = 0.2$, $C_U = 1$ and $C_J = 0.5$.

As shown in Fig. 3(a), the BER of the signals sent by OBUs in the UAV-aided VANET decreases with time, e.g., from 3.3% at the beginning of the game to 0.13% after 1500 time slots, about 80% lower than that of the Q-learning based strategy. This is because the UAV learns and adjusts its relay strategy accordingly, and the PHC-based relay strategy increases the randomness compared to Q-learning based strategy at the beginning of exploration.

As shown in Fig. 3(b), the utility of the UAV increases with time, and the PHC-based relay strategy has a higher utility than the Q-learning based relay strategy. For example, the utility of the UAV increases by about 53% after 1500 time slots at convergence, about 12% higher than that of the Q-learning strategy. In addition, the convergent speed of the hotbooting PHC increases by 75% compared with the standard PHC strategy, because the hotbooting technique can formulate the emulated experiences to reduce the exploration trials and thus significantly improve the learning speed and jamming resistance efficiency.

VII. Conclusion

In this paper, we have formulated a UAV-aided VANET transmission game to address smart jamming, and derived the NEs of the anti-jamming UAV relay game to disclose the impacts of the transmission cost and the radio channel condition on the jamming resistance of VANET. A hotbooting PHC-based anti-jamming relay strategy has been proposed for a UAV without the knowledge of VANET network model and jamming model. Simulation results have shown that the proposed relay strategy can efficiently improve the anti-jamming transmission performance. For example, the BER of the signal sent by OBUs in the VANET with the PHC-based relay strategy decreases by 80% and the utility of the UAV increases by 12% compared with the Q-learning based relay strategy.

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] 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.

[4] 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.

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

[6] X. Du and H. H. Chen, "Security in wireless sensor networks," *IEEE Wireless Comm. Magazine*, vol. 15, no. 4, pp. 60–66, Aug. 2008.

[7] Y. Xiao, H. H. Chen, X. Du, and M. Guizani, "Stream-based cipher feedback mode in wireless error channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 662–666, Feb. 2009.

[8] S. Liang and X. Du, "Permission-combination-based scheme for android mobile malware detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, Australia, Jun. 2014.

[9] L. Xiao, N. B. Mandayam, and H. V. Poor, "Prospect theoretic analysis of energy exchange among microgrids," *IEEE Trans. Smart Grids*, vol. 1, no. 2, pp. 63–72, Jun. 2015.

[10] 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, p-p. 2068–2085, Jul. 2011.

[11] 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.

[12] 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.

[13] 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.

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

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

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