Joint D2D Cooperative Relaying and Friendly Jamming Selection for Physical Layer Security

Yijie Luo[1(✉)], Yang Yang[1], Yanlei Duan[2], and Zhengju Yang[2]

[1] Army Engineering University of PLA, Nanjing, China
yijieluo@sina.com, sheep_1009@163.com
[2] Troop of PLA, Kunming, China
duanyanlei2008@163.com, yangzhengjul001@126.com

Abstract. D2D communications are emerging technologies to improve spectrum efficiency, energy efficiency as well as physical layer security of cellular networks. In most research, D2D users, considered as friendly jammers, can improve the information security of cellular networks. D2D users can also work as cooperative relays between the eNB and the cellular user (CU) to increase the transmission rate and improve the security capacity simultaneously. Considering there exists an active eavesdropper in the cellular network, which can attack as a passive eavesdropper or an active jammer, joint D2D cooperative relaying and friendly jamming selection can enhance the secrecy achievable rate or the transmission rate of CU. We formulate a Stackelberg game between different intelligent agents, and derive the mixed-strategy equilibrium (MSE) via a hierarchical learning algorithm based on Q-learning. Simulation results show that the strategic selections of D2D users and the active eavesdropper are convergent, and the proposed algorithm has a better performance than the random selection method.

Keywords: D2D communications · Cooperative relaying · Friendly jamming
Active eavesdropper · Stackelberg game · Q-learning

1 Introduction

D2D communications, owning to its potential ability to realize low-latency and high-data-rate communications and bring higher spectrum and energy efficiency, are considered as a disruptive technology direction of 5G mobile communication systems [1, 2]. Most of researches of D2D communications focus on resource allocation and interference avoidance [3, 4]. While introducing D2D communications to cellular networks, security problem becomes more and more important. In [5], secrecy energy efficiency of the cellular user and the D2D user was improved by the proposed power control algorithm based on Stackelberg game. In [6, 7], considering power control, access control and D2D pair selection, a joint mechanism was proposed to enhance secrecy performance. Furthermore, the cooperation among cellular users and D2D users was formulated as a coalitional game, and both social welfare and system secrecy rate were improved via proposed algorithms. In these works, D2D users were always considered to be friendly jammers to enhance system achievable secrecy rate by

L. Meng and Y. Zhang (Eds.): MLICOM 2018, LNICST 251, pp. 115–126, 2018.
https://doi.org/10.1007/978-3-030-00557-3_12

deteriorating the wiretap channel. There were also some works considering D2D users to be cooperative relays in cellular networks, while not in the physical layer security aspect. In [8], cooperation schemes in the form of relaying or jamming between cellular and D2D users was discussed, and three cooperative frameworks were proposed and compared. Considering the channel estimation error and the interference between different D2D users, a joint beamforming design of base station and D2D relay user was proposed in [9]. In [10], joint mode selection, resource assignment and power allocation scheme of D2D users were studied and the joint relay and jamming scheme outperformed conventional D2D directly or relay communication schemes.

In fact, cooperative relaying can enhance achievable secrecy rate by improving the transmission rate of legitimate users. Joint cooperative relaying and friendly jamming selection is widely applied to improve physical layer secrecy performance in all kinds of wireless networks [11–14]. In [11], the reliable-and-secure connection probability (RSCP) and the reliability-security ratio (RSR) were introduced and analyzed in different relaying and jamming cooperation schemes with channel state information (CSI) feedback delays. In [12], the ergodic secrecy rate (ESR) was maximized and joint power allocation and relay selection scheme was presented. In [13], considering intermediate nodes working as relays or jammers, two relay and jammer selection approaches were proposed to minimize secrecy outage. In [14], the scheme that source node communicating with destination node securely via cooperative relay and cooperative jammer selection was considered, and a particle swarm optimization approach was proposed to enhance overall secrecy achievable rate. Whereas few works on D2D relay-assisted secrecy cellular communications is under studied. So we consider introducing joint cooperative relaying and friendly jamming selection strategies to D2D communications for physical layer security improvement. In our former work [15], we consider single D2D user to work as a cooperative relay or a friendly jammer to enhance secrecy achievable rate of the cellular user, a non-cooperative game between the D2D user and the active eavesdropper was formulated and the mixed-strategy Nash equilibrium (MSNE) was derived via the fictitious play-based algorithm [16].

In this paper, we go further to consider multiple D2D users underlaying cellular networks where there exists an intelligent attacker (active eavesdropper), who has the dual ability of either passively eavesdropping or actively jamming cellular links. Due to the mutually opposite interest of the legitimate user and the attacker, a Stackelberg game is formulated between them. To be specific, the legitimate user is modeled as the leader and the active eavesdropper is the follower. The legitimate user firstly selects a "best" D2D user to work as a cooperative relay and leaving other D2D users to work as friendly jammers. Then the active eavesdropper selects passively eavesdropping or actively jamming as follows. We analyze the existence of MSE of the Stackelberg secrecy game and achieve a MSE using a hierarchical Q-learning algorithm with which the legitimate user and the attacker can update their strategies. The contributions of the paper are as follows: (1) compared with our former work, we consider multiple D2D users cooperation rather than single D2D user access, (2) we propose a joint friendly jammer and cooperative relay selection scheme to retain the diversity gain, enhance physical layer security and increase transmission opportunity of D2D users to the utmost, (3) The proposed hierarchical learning algorithm based on Q-learning significantly outperforms the random selection method.

The rest of this paper is organized as follows. In Sect. 2, the system model is described and the secrecy rate under different schemes is analyzed. In Sect. 3, utility functions of cellular networks and the active eavesdropper are designed and a Stackelberg game between them is formulated. Then the Q-learning based algorithm is proposed to find the MSE and simulation results are analyzed and discussed in Sect. 4. And conclusions are drawn in Sect. 5 finally.

2 System Model and Problem Formation

We consider a single cell scenario, where there is an evolved Node B (eNB), a cellular user, N D2D users and an active eavesdropper, which are equipped with a single antenna and operate in a half-duplex mode, as shown in Fig. 1. The eNB, the cellular user and the active eavesdropper are denoted as B, C and A respectively. All the N D2D users form the set of $\mathcal{N}$, one of D2D users is denoted as $n_i \in \mathcal{N}$. Furthermore, we assume that the eNB can establish the direct cellular link to the cellular user, and select only one D2D transmitter to relay confidential data to the destination cellular user, while other $N-1$ D2D users transmit their own data through direct D2D channels in the underlay way, respectively. And assumed that the attacker can work in two modes: (1) passively overhearing the confidential information transmitted from the eNB and the relay D2D transmitter; (2) actively jamming signals received by cellular user.

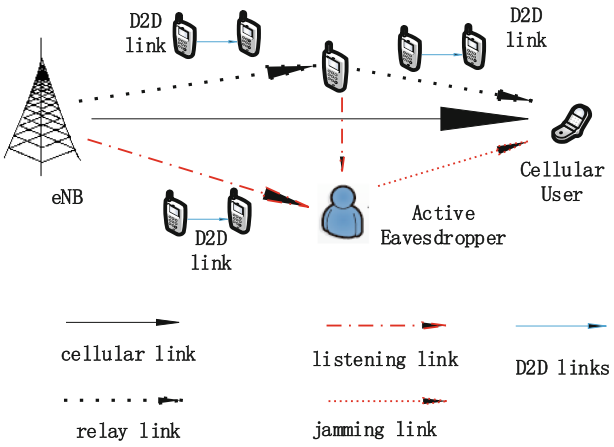

Fig. 1. System model

Suppose that the transmission power of the eNB, the jamming power of the active eavesdropper and the transmission (or relaying) power of the *ith* D2D transmitter are denoted by P_B, P_J and P_i, respectively. The channel gain between the eNB and the cellular user, between the eNB and the active eavesdropper, and between the cellular user and the active eavesdropper, are denoted by h_{BC}, h_{BA} and h_{AC}, respectively. The channel gain between the eNB and the *ith* D2D transmitter, between the active

eavesdropper and the *ith* D2D transmitter, between the *ith* D2D transmitter and the cellular user, between the *ith* D2D transmitter and the *jth* D2D receiver are denoted by h_{Bi}, h_{Ai}, h_{iC} and h_{ij}, respectively. And the background additive white Gaussian noise is denoted by N_0.

2.1 Secrecy Achievable Rate of Cellular User When Passively Eavesdropped

Because the active eavesdropper works in a half-duplex mode, it can't eavesdrop and jam simultaneously. Therefore, we first consider it works as a passively eavesdropper. It is assumed that the eNB can transmit the confidential information not only through the direct cellular link to the cellular user, but also through the selected D2D transmitter to relay. In the first phase, the confidential information is transmitted from the eNB to the D2D transmitter (suppose that it is the *ith* D2D transmitter); while in the second phase, the selected *ith* transmitter relays the received information to the cellular user. Meanwhile the confidential information transmission is performed through the direct cellular link during the two phases. Hence the signal-to-interference-plus-noise-power-ratio (SINR) for the direct cellular link and the direct eavesdropping link are expressed as

$$\xi_1 = \frac{P_B|h_{BC}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{jC}|^2 + N_0} \tag{1}$$

and

$$\phi_1 = \frac{P_B|h_{BA}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{jA}|^2 + N_0} \tag{2}$$

respectively.

If the eNB selects the *ith* D2D transmitter to be the relay node, and the relay node employs Amplify and Forward (AF) protocol to forward the message. It is also assumed that the eavesdropper is mainly interested in the confidential information, so it wiretaps the cellular link and relaying link where the confidential information going through. Then the SINR for the relay cellular link and the relay eavesdropping link are

$$\xi_2 = \frac{\gamma_{1E} \cdot \gamma_{2E}}{\gamma_{1E} + \gamma_{2E} + 1} \tag{3}$$

where $\gamma_{1E} = \frac{P_B|h_{Bi}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{ij}|^2 + N_0}$, $\gamma_{2E} = \frac{P_i|h_{iC}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{jC}|^2 + N_0}$,

and

$$\phi_2 = \frac{\psi_1 \cdot \psi_2}{\psi_1 + \psi_2 + 1} \tag{4}$$

where

$\psi_1 = \frac{P_B|h_{Bi}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{ij}|^2 + N_0}, \psi_2 = \frac{P_i|h_{iA}|^2}{\sum_{j=1,j\neq i}^{N-1} P_j|h_{jA}|^2 + N_0}$

respectively. Therefore the achievable rate of CU and the wiretap rate of the eavesdropper are expressed as

$$R_R(i,E) = \frac{1}{2}\log_2(1 + \xi_1 + \xi_2) \tag{5}$$

and

$$R_E(i,E) = \frac{1}{2}\log_2(1 + \phi_1 + \phi_2) \tag{6}$$

respectively. Then the achievable secrecy rate of CU is expressed as

$$R(i,E) = [R_R(i,E) - R_E(i,E)]^+ \tag{7}$$

where $[\cdot]^+ = \max\{\cdot, 0\}$.

2.2 Transmission Rate of Cellular User When Actively Jammed

If the attacker chooses the active jamming mode, the transmission signals of D2D users are not friendly at all but harmful to the cellular user for introducing more extra interference. While compared with working as cooperative relays, D2D users prefer to work as friendly jammers to transmit their own data. Therefore, we consider the scheme that the eNB selects a "best" D2D transmitter to relay data to the cellular user, and other D2D pairs transmit their own data. It can retain the diversity gain, eliminate the interference to the cellular link and increase transmission opportunity of D2D users to the utmost.

Under this scheme, the SINR of CU for the direct cellular link when jammed by the active eavesdropper is

$$\omega_1 = \frac{P_B|h_{BC}|^2}{P_J|h_{AC}|^2 + \sum_{j=1,j\neq i}^{N-1} P_j|h_{jC}|^2 + N_0} \tag{8}$$

The SINR for the relaying cellular link when jammed by the active eavesdropper is

$$\omega_2 = \frac{\gamma_{1J} \cdot \gamma_{2J}}{\gamma_{1J} + \gamma_{2J} + 1} \tag{9}$$

where

$$\gamma_{1J} = \frac{P_{Bi}|h_{Bi}|^2}{P_J|h_{Ai}|^2 + \sum_{j=1,j\neq i}^{N-1} P_j\left|h_{ij}\right|^2 + N_0}, \ \gamma_{2J} = \frac{P_i|h_{iC}|^2}{P_J|h_{AC}|^2 + \sum_{j=1,j\neq i}^{N-1} P_j\left|h_{jC}\right|^2 + N_0}.$$

Hence when the active eavesdropper selects actively jamming, the achievable rate of the CU is expressed as

$$R(i,J) = \frac{1}{2}\log_2(1+\omega_1+\omega_2) \tag{10}$$

Let $a = E, J$, then the utility function of the legitimate user is

$$U(i,a) = R(i,a) \tag{11}$$

While the utility function of the active eavesdropper is

$$U_A(i,a) = \begin{cases} -R(i,E), a = E \\ -R(i,J) - c_J P_J, a = J \end{cases} \tag{12}$$

where c_J represents a cost factor on the power level used by the active eavesdropper when it chooses actively jamming.

3 Stackelberg Game Formulation

In this section, a Stackelberg game is formulated to characterize the interaction between legitimate users and the active eavesdropper. Specifically, the legitimate user first selects the "best" D2D transmitter to maximize its utility function, which acts as the leader, whereas the active eavesdropper selects attacking modes subsequently to minimize its jamming cost, which is the follower. The strategy of the legitimate user is the probability of selecting which D2D user to relay its message to the cellular user and leaving other D2D users to transmit their own data to be friendly jammers. The probability of selecting the *ith* D2D user is denoted by p_{n_i}, and let $\mathbf{P}_n = [p_{n_1}, p_{n_2}, \ldots, p_{n_N}]$ be a mixed strategy of the legitimate user in the set of all feasible strategies $\mathcal{P}_n := \{p_{n_i} \in \mathbb{R}_+ : \sum_{n_i \in \mathcal{N}} p_{n_i} = 1\}$.While the active eavesdropper's strategy is the probability of selecting passively eavesdropping or actively jamming. Let p_E and p_J be the eavesdropping probability and jamming probability respectively, $\mathbf{P}_A = [p_E, p_J]$ is an admissible mixed strategy of the active eavesdropper and $\mathcal{P}_A$ is the set of admissible mixed strategies, defined by $\mathcal{P}_A := \{p_E, p_J \in [0,1]^2 : p_E + p_J = 1\}$. Then average utilities of the legitimate user and the active eavesdropper are expressed as:

$$\overline{U}(\mathbf{P}_n, \mathbf{P}_A) = \mathbb{E}_{\mathbf{P}_n, \mathbf{P}_A}(U(i,a)) \tag{13}$$

$$\overline{U_A}(\mathbf{p}_n, \mathbf{P}_A) = \mathbb{E}_{\mathbf{P}_n, \mathbf{P}_A}(U_A(i,a)) \tag{14}$$

Proposition 1. Let $\mathcal{G} = \{\{B, A\}, \mathcal{N}, \{a\}, \overline{U}, \overline{U_A}\}$ be the game described in Sect. 3, the Stackelberg game admits a MSE $(\mathbf{P}_n^*, \mathbf{P}_A^*)$, which satisfies the following set of inequalities:

$$\overline{U}(\mathbf{P}_n^*, \mathbf{P}_A^*) \geq \overline{U}(\mathbf{P}_n^*, \mathbf{P}_A), \forall \mathbf{P}_A \in \mathcal{P}_A \tag{15}$$

$$\overline{U_A}(\mathbf{P}_n^*, \mathbf{P}_A^*) \geq \overline{U_A}(\mathbf{P}_n, \mathbf{P}_A^*), \forall \mathbf{P}_n \in \mathcal{P}_n, n_i \in \mathcal{N} \tag{16}$$

Proof. Since every finite strategy game has a MSE [17], we can achieve the above outcome on existence of the MSE of the Stackerlberg game. ■

The MSE of the game defines a state in which no player, including the legitimate user and the active eavesdropper, has an incentive to change its strategies.

4 Algorithm Description

In this section, we will study the MSE of the proposed secrecy game. Based on Q-learning algorithm [18, 19], we propose a hierarchical algorithm to find an MSE of the game $\mathcal{G}$. Firstly, the legitimate user makes its strategic decisions according to its policy $\mathbf{P}_n(k)$, which is the mixed strategy of D2D users in the kth epoch. And then the active eavesdropper updates its Q value in the tth iteration as follows:

$$Q_{A,m}(t+1) = (1 - \kappa_A^t) Q_{A,m}(t) + \kappa_A^t U_A(t), \tag{17}$$

where $\kappa_A^t \in [0, 1)$ is the learning rate of the active eavesdropper, and $\sum_{t=0}^{\infty} \kappa_A^t = \infty$, $\sum_{t=0}^{\infty} (\kappa_A^t)^2 < \infty$. The attacker's strategy is updated according to:

$$p_{A,m}(t+1) = \frac{\exp[Q_{A,m}(t)/\tau_0]}{\sum_{r \in \mathcal{A}} \exp[Q_{A,r}(t)/\tau_0]}, \tag{18}$$

where τ_0 controls the tradeoff of exploration-exploitation.

Then the Q value of the legitimate user is updated as:

$$Q_{B,n}(k+1) = (1 - \kappa_B^k) Q_{B,n}(k) + \kappa_B^k U(k), \tag{19}$$

where $\kappa_B^k \in [0, 1)$ is the learning rate of the legitimate user, and $\sum_{k=0}^{\infty} \kappa_B^k = \infty$, $\sum_{k=0}^{\infty} (\kappa_B^k)^2 < \infty$. Then the probability of selecting the relay is updated according to

$$p_{B,n}(k+1) = \frac{\exp[Q_{B,n}(k)/\tau_0]}{\sum_{w \in \mathcal{N}} \exp[Q_{B,w}(k)/\tau_0]} \tag{20}$$

Theorem 1. The proposed hierarchical learning algorithm can discover a MSE strategy of the secrecy game.

Proof. For brevity, the convergence of the proposed hierarchical algorithm can be found in [20] and it is a MSE of the secrecy game. ■

Then this algorithm is summarized in Table 1.

Table 1 .

TABLE I. **ALGORITHM 1:** PROPOSED HIERARCHICAL LEARNING ALGORITHM BASED ON Q-LEARNING

1: **initialization:**
2: **set** $t=0$, $k=0$, Q values of the eNB and the attacker
3: **set** initial selecting probabilities $p_{n_i}=\frac{1}{N}, n_i \in N$
4: **set** initial attacking probabilities $p_E = p_J = \frac{1}{2}$
5: **end initialization**
6: According to the policy $\mathbf{P}_B(k)$ of the legitimate users in the kth epoch, the eNB selects a D2D transmitter from the D2D users' set $\mathcal{N}$.
7: The active eavesdropper's learning process.
8: **innerloop**
(1) According to the policy $\mathbf{P}_A(t)$ in the tth slot, the active eavesdropper selects its attacking strategy from the set $\mathcal{A}$.
(2) The active eavesdropper calculates its utility function $U_A(t)$ in the tth slot.
(3) According to (17) , the active eavesdropper updates its Q values and according to (18), it updates its attacking policy.
9: **set** t=t+1, and until the stopping criterion hold.
10: **end inner loop**
11: The legitimate user calculates its utility function $U(k)$ in the kth epoch.
12: According to (19), the legitimate user updates its Q values and according to (20), it updates its relay selection probability.
13: Go to 6, and until the stopping criterion hold.

5 Simulation and Numerical Analysis

For our simulations, we consider a D2D underlay cellular network composed of a square area of 1 km * 1 km with the eNB located at the center and the cellular user and all the D2D users are randomly located on the square area of 0.5 km * 0.5 km centered

by the eNB, while the active eavesdropper is randomly located between the square area of 0.5 km * 0.5 km and 1 km * 1 km. In these simulations, a path loss model is adopted, and the path loss exponent is set as $\alpha = 3$. The transmit power of the eNB and D2D users are set as $P_B = 1\,\mathrm{W}$,$P_i = 100\,\mathrm{mW}$, and the jamming power of the active eavesdropper is set as $P_J = 100\,\mathrm{mW}$, respectively. The jamming cost is set as $c_J = 1$, and the noise level is set as $N_0 = 10^{-10}\,\mathrm{W}$.

In Fig. 2, the update of the active eavesdropper's attacking mode selection probability in the first epoch is presented, while in Fig. 3, the convergent process of the D2D relay selection over epoch numbers is showed. From these figures, we have found that the selection probability of the legitimate user (or the active eavesdropper) converges to a stationary mixed strategy very soon. In Fig. 4, it is compared with the random selection algorithm (RSA) to evaluate the proposed hierarchical learning algorithm based on Q-learning (HLA). In RSA, the legitimate user and the active eavesdropper randomly select their actions from their strategy sets at each time. It clearly shows that the proposed HLA presents a significant better performance gain than RSA at all sizes of D2D users. And it also shows that the average expected utility of CU decreases as the number of D2D users increases, and the gaps between different algorithms will be narrowed. That is because the more D2D users, the larger interference they induce to the CU, and they lower the transmission rate of the CU dramatically.

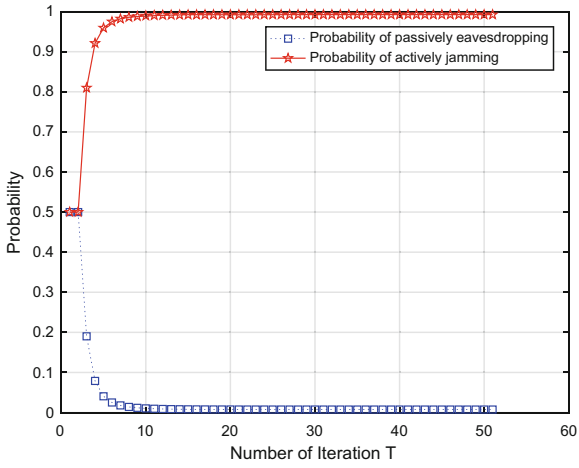

Fig. 2. The convergent process of the active eavesdropper's attacking mode selection.

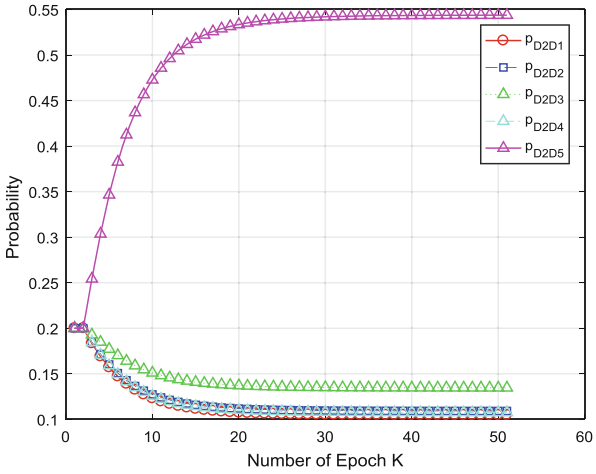

Fig. 3. The convergent process of D2D relay selection.

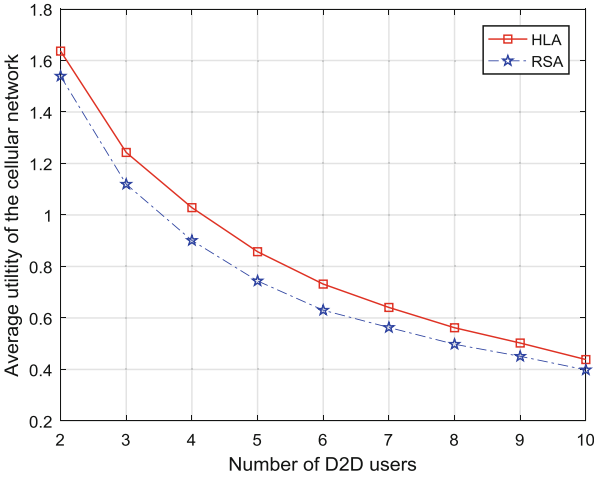

Fig. 4. Average utility function of the cellular network vs. number of D2D users

6 Conclusions

In this paper, we have formulated a Stackelberg game between the legitimate user and the active eavesdropper. Firstly, the eNB works as the leader, select a "best" D2D user to work as the cooperative relay and others as friendly jammers to improve physical layer security of cellular networks, then the active eavesdropper, working as the follower, selects the passive eavesdropping or the active jamming mode to lower the overall secrecy rate or transmission rate of cellular networks. To solve this problem, we have employed a hierarchical learning algorithm based on Q-learning using which the

legitimate user and the active eavesdropper can reach a MSE through a small number of iterations. The results have shown that the proposed algorithm enables the legitimate user to significantly improve its average expected utility either the achievable secrecy rate of the cellular link when eavesdropped or the transmission rate when jammed.

References

1. Federico, B., Robert, W.H., Angel, L., Thomas, L.M., Petar, P.: Five disruptive technology directions for 5G. IEEE Commun. Mag. **52**(02), 74–80 (2014)
2. Shaoyu, L., Chunche, C., Fanmin, T., Tienchen, H.: 3GPP device-to-device communications for beyond 4G cellular networks. IEEE Commun. Mag. **54**(03), 28–35 (2016)
3. Zhenyu, Z., Mianxiong, D., Kaoru, O., Jun, W., Takuro, S.: Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications. IEEE Commun. Lett. **3**(05), 485–488 (2014)
4. Lili, W., Rose, Q.H., Yi, Q., Geng, W.: Enable device-to-device communications underlaying cellular networks: challenges and research aspects. IEEE Commun. Mag. **52** (06), 90–96 (2014)
5. Wanbing, H., Wei Z., Wei B., Yueming C., Xinrong G., Junyue Q.: Improving physical layer security in underlay D2D communication via stackelberg game based power control. In: 2016 IEEE International Conference on Computer, Information and Telecommunication Systems (CITS), pp. 1–5. IEEE Press, Kunming (2016)
6. Rongqing, Z., Xiang, C., Liuqing, Y.: Joint power and access control for physical layer security in d2d communications underlaying cellular networks. In: 2016 IEEE International Conference on Communications (ICC), pp. 1–6, IEEE Press, Kuala Lumpur (2016)
7. Rongqing, Z., Xiang, C., Liuqing, Y.: Cooperation via spectrum sharing for physical layer security in device-to-device communications underlaying cellular networks. IEEE Trans. Wirel. Commun. **15**(8), 5651–5663 (2016)
8. Yang, C., Tao, J., Conggang, W.: Cooperative device-to-device communications in cellular networks. IEEE Wirel. Commun. **22**(3), 124–129 (2015)
9. Yi, Q., Ming, D., Meng, Z., Hui, Y., Hanwen, L.: Relaying robust beamforming for device-to-device communication with channel uncertainty. IEEE Commun. Lett. **18**(10), 1859–1862 (2014)
10. Tuong, D.H., Long, B.L., Tho, L.N.: Joint mode selection and resource allocation for relay-based d2d communications. IEEE Commun. Lett. **21**(2), 398–401 (2017)
11. Lei, W., Yueming, C., Yulong, Z., Weiwei, Y., Lajos, H.: Joint relay and jammer selection improves the physical layer security in the face of CSI feedback delays. IEEE Trans. Veh. Technol. **65**(8), 6259–6274 (2015)
12. Chao, W., Huiming, W., Xia, X.: Hybrid opportunistic relaying and jamming with power allocation for secure cooperative networks. IEEE Trans. Wirel. Commun. **14**(2), 589–605 (2015)
13. Hui, H., Lee, S., Guobing, L., Junli, L.: Secure relay and jammer selection for physical layer security. IEEE Signal Process. Lett. **22**(8), 1147–1151 (2015)
14. Ning, Z., Nan, C., Ning, L., Xiang, Z., Jon, W.M., Xuemin (Sherman), S.: Partner selection and incentive mechanism for physical layer security. IEEE Trans. Wirel. Commun. **14**(8), 4265–4276 (2015)
15. Yijie, L., Yang, Y., Cui, L.: Research on physical layer security in D2D enabled cellular networks with an active eavesdropper. Signal Proccesing (accepted)

16. Fudenberg, D., Levine, D.K.: The Theory of Learning in Games. MIT Press, Cambridge (1998)
17. Han, Z., et al.: Game Theory in Wireless and Communication Networks. Cambridge University Press, Cambridge (2012)
18. Watkins, C.J.C.H., Dayan, P.: Q-learning. Mach. Learn. **8**, 279–292 (1992)
19. Luliang, J., Fuqiang, Y., Youming, S., Yuhua, X., Shuo, F., Alagan, A.: A hierarchical learning solution for antijamming stackelberg game with discrete power strategies. IEEE Wirel. Commun. Lett. **6**(6), 818–821 (2017)
20. Sastry, P.S., Phansalkar, V.V., Thathachar, M.: Decentralized learning of nash equilibria in multi-person stochastic games with incomplete information. IEEE Trans. Syst. Man Cybern. **24**(5), 769–777 (1994)