Game Theoretic Study on Blockchain Based Secure Edge Networks

Dongjin Xu*, Liang Xiao*, Limin Sun†, Min Lei‡
* Dept. of Communication Engineering, Xiamen University, China. Email: lxiao@xmu.edu.cn
† Beijing Key Laboratory of IOT Information Security Technology, Institute of Information Engineering, CAS, China.
Email:sunlimin@iie.ac.cn
‡ Information Security Center, Beijing University of Posts and Telecommunications, China
Email:leimin@bupt.edu.cn

***Abstract*—Blockchain has been applied to study data privacy and network security recently. In this paper, we propose a punishment mechanism based on the action record on the blockchain to suppress the attack motivation of the edge servers and the mobile devices in the edge network. The interactions between a mobile device and an edge server are formulated as a blockchain security game, in which both of them are able to launch a wide range of attacks, and decide to whether or not to attack each other. A Q-learning based security scheme is proposed for the device without being aware of the network condition in the dynamic blockchain security game. Simulation results show that the proposed security scheme prevents adversary behaviors of the edge server and improves the utility of the mobile device, compared with the benchmark greedy strategy.**

***Index Terms*—Blockchain, reinforcement learning, game theory, network security.**

I. Introduction

The real-time mobile applications such as augmented reality and gaming applications require the low-delay mobile services to provide seamless end-user interaction. Mobile devices can offload their application execution to the cloud servers to improve user experience in terms of faster processing speed, longer battery lifetime and more powerful security services. However, remote cloud servers can incur long delays and poor connectivity over the Internet. Moving cloud servers to the edge of the network reduces the delays while preserving the essential benefits of a high-performance cloud, which is intended to enable a range of latency sensitive mobile applications [1]. Since the edge servers and the mobile devices are rational and thus naturally selfish, they make any efforts to maximize their payoffs including launching a range of attacks such as denial-of-service (DoS) attacks.

As an emerging decentralized architecture, blockchain has recently attracted intensive attention from governments, financial institutions and capital markets due to its decentralization, collective maintenance, programmability and security [2]. Blockchain was first introduced as a decentralized transparent ledger that records transactions across all network nodes without a third-party intermediary [3]. In the blockchain network, each node performing recording and relaying transactions has a copy of the chain with complete information about the transactions that cannot be modified retroactively. Therefore, blockchain can provide insights into network security and data privacy [4]–[6].

Game theory has been used to study the defense mechanism against attacks in the networks [7], [8]. In this paper, we formulate a blockchain security game, in which the mobile device and the edge server are able to launch a range of attacks, and propose a blockchain based punishment mechanism to suppress their attack motivation. The Nash equilibria (NEs) of the game are derived and the conditions that each NE exists are provided to disclose how the punishment mechanism impacts the attack rates of the mobile device and the edge server.

As the decision making process of the device in the dynamic game can be viewed as a Markov decision process (MDP), reinforcement learning (RL) such as Q-learning can achieve the optimal strategy via trials-and-errors if the game is long enough [9], [10]. Therefore, we propose a Q-learning based security scheme for the device to prevent adversary behaviors of the server without the knowledge of the network condition in the dynamic game. Simulation results show that the proposed security strategy can efficiently reduce the attack rate of the server and increase the utility of the device. The contributions of this work can be summarized as:

- We formulate a blockchain security game to investigate the punishment mechanism based on the action history recorded on the blockchain. The NEs of the game and the conditions under which the NEs exist are provided.
- We propose a Q-learning based security scheme to prevent adversary behaviors of the server without the knowledge of the network condition. Simulation results show that this scheme achieves a lower attack rate of the server and a higher utility compared with a greedy strategy.

The remainder of the paper is organized as follows. We review related work in Section II and present the system model in Section III. We present the blockchain security game and provide the NE of the game in Section IV. We propose a Q-learning based security scheme in the dynamic game in Section V, and provide simulation results in Section VI. Conclusions are drawn in Section VII.

II. Related Work

Blockchain has been applied to study data privacy and network security. A personal data management system as de-

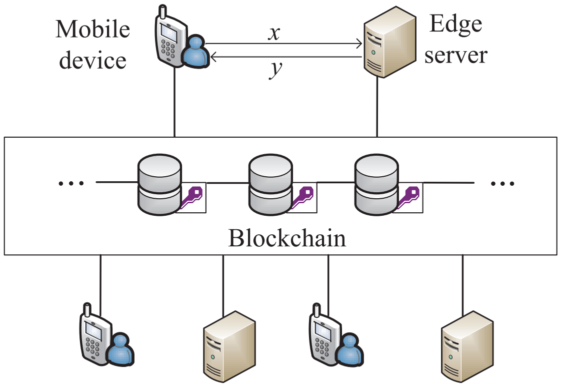


Fig. 1: Illustration of a blockchain based edge network.

signed in [3] uses blockchain as an access-control manager to provide personalized services for mobile users. A blockchain model of cryptography as formulated in [4] protects transactional privacy in the decentralized smart contract system. A cloud-centric IoT system as proposed in [5] uses micro services that can engage other restful services to process requests from the edge devices to reduce the network latency, and applies blockchain to provide secure and persistent data storage. A blockchain-based smart home framework as formulated in [6] reduces energy consumption and data packet processing overhead of the Internet of Things (IoT) devices. An intelligent transportation system as proposed in [11] utilizes blockchain to provide realtime ride-sharing services for the vehicles.

Game theory has provided insights into network security against attacks. A security game as formulated in [12] investigates the defense mechanism against co-resident attacks to maximize the attack cost and minimize the probability of achieving co-residence. An invalid signature identification game as investigated in [13] enables mobile nodes to detect attacks with the optimal delay in wireless mobile networks. A Q-learning based mobile offloading strategy as proposed in [14] reduces the attack rate of the smart attackers in the dynamic offloading game. The two-layer advance persistent threat defense game as formulated in [15] studies the joint threats from an attacker and insiders in the cyber system. A repeated Bayesian game as modeled in [16] changes the configurations of the web applications such as the coding language against attacks.

III. System Model

We consider a blockchain based edge network consisting of M edge servers (S) and N mobile devices (D) that each have a decentralized record on the interactions among them as shown in Fig. 1. A mobile device sends a request to an edge server at time k to obtain real-time service or launches attacks against the server for illegal security gains. We assume the device is able to launch a wide range of attacks at each time, such as DoS attacks, zero-day attack, and password-based attacks [17]. The attacks are split into different levels according to their impacts on the network performance and their attack costs.

TABLE I: Action Set of the Mobile Device.

Action ID	Physical action
0	Request services
1	Level-1 attack, e.g., fishing
2	Level-2 attack, e.g., password-based attack
...	...
$L-1$	Level-$(L-1)$ attack, e.g., DoS
L	Level-L attack, e.g., zero-day

We assume Level-i attack is more dangerous to the network than Level-j attack, with $i \geq j$. For example, a zero-day attack can be labelled with a higher level than a DoS attack for the higher security risk. The action of the mobile device at time k is denoted by $x \in \{i\}_{0 \leq i \leq L}$. As shown in Table 1, the device sends a service request to the server if $x = 0$, and launches Level-x attack if $1 \leq x \leq L$.

The server chooses to perform the request from the device or attack it at time k, denoted by $y \in \{j\}_{0 \leq j \leq L}$. Similarly, we assume the server is able to launch a variety of attacks that are split into different levels according to their danger gradations. The server follows the request from the device if $y = 0$, and launches Level-y attack if $1 \leq y \leq L$.

Based on the blockchain network, the system does not require a centralized process or a central unit. Each node in the network checks the actions of the others, updates the action records, and shares the new records over the network. The records that are allocated to each node according to its action history determine its probability to receive the network services or node cooperation in the network. The aggressive behaviors recorded on the blockchain are punished by the other nodes in the network to improve the security performance of the edge network.

IV. Blockchain Security Game

The interaction between an edge server and a mobile device can be formulated as a blockchain security game. As shown in Fig. 1, the mobile device chooses to send a service request to the server or launch attacks against the server, $x \in \mathbf{x} = \{i\}_{0 \leq i \leq L}$. The edge server chooses to follow the request or launch attacks against the device, $y \in \mathbf{y} = \{j\}_{0 \leq j \leq L}$.

Let G_x^y denote the direct benefit of the mobile device choosing x if the edge server takes the action y at time k, which is the gain from the service if $x = y = 0$, the security loss if $y > x = 0$, the illegal gain minus the attack cost if $x > y = 0$, and the illegal gain minus the attack cost and the security loss otherwise. Similarly, the benefit of the server choosing y if the mobile device takes the action x at time k is denoted by C_x^y, which represents the service cost if $x = y = 0$, the sum of the service cost and the security loss if $x > y = 0$, the illegal gain minus the attack cost if $y > x = 0$, and the illegal gain minus the attack cost and the security loss otherwise. The payment to the edge server that follows the request at time k is denoted by $R^{(k)}$. The punishment factor

TABLE II: Summary of Symbols and Notation.

Notation	**Definition**
x/**y**	Action set of the mobile device/edge server
G_x^y	Direct benefit of the mobile device choosing x if the edge server takes the action y
C_x^y	Direct benefit of the edge server choosing y if the mobile device takes the action x
$R^{(k)}$	Payment to the edge server at time k
$\beta^{(k)}$	Punishment factor at time k
$u_D^{(k)}/u_S^{(k)}$	Utility of the mobile device/edge server at time k

at time k denoted by $\beta^{(k)}$ represents the weight of the action record on the blockchain.

The utility of the mobile device at time k, denoted by $u_D^{(k)}$, depends on the direct gain, the payment to the server, and the punishment for attacks, and is given by

$$u_D^{(k)}(x,y) = G_x^y - \mathrm{I}(x=y=0)R^{(k)} - x\beta^{(k)}. \qquad (1)$$

The utility of the server at time k denoted by $u_S^{(k)}$ is given by

$$u_S^{(k)}(x,y) = C_x^y + \mathrm{I}(x=y=0)R^{(k)} - y\beta^{(k)}. \qquad (2)$$

The time index k in the superscript is omitted if no confusion occurs. Table 2 summarizes the notation used in the paper.

An NE of the blockchain security game, denoted by (x^*, y^*), consists of the best responses of the mobile device and the edge server if the opponent uses the NE strategy. By definition, we have

$$x^* = \arg\max_{x\in\mathbf{x}} u_D(x, y^*) \qquad (3)$$

$$y^* = \arg\max_{y\in\mathbf{y}} u_S(x^*, y). \qquad (4)$$

Theorem 1. *The blockchain security game has an NE* $(x^*, y^*) = (0, 0)$*, if*

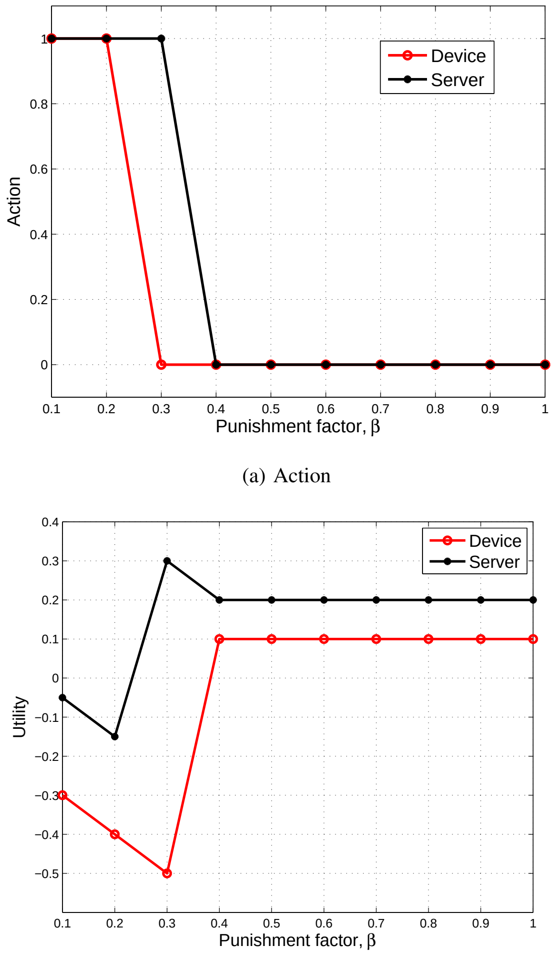
$$G_0^0 - R \geq \max_{1\leq i\leq L} \left(G_i^0 - i\beta\right) \qquad (5)$$

$$C_0^0 + R \geq \max_{1\leq i\leq L} \left(C_0^i - i\beta\right). \qquad (6)$$

Proof: By (1), if (5) holds, ∀ $1 \leq x \leq L$, we have

$$u_D(0,0) = G_0^0 - R \geq G_x^0 - x\beta = u_D(x,0). \qquad (7)$$

Thus, (5) holds for $(x^*, y^*) = (0, 0)$. Similarly, we have (6) holds for $(x^*, y^*) = (0, 0)$, indicating that $(0, 0)$ is an NE of the game. ■



(a) Action

(b) Utility

Fig. 2: Performance of the blockchain security game at the NE, with $R = 0.6$ and $L = 1$.

Theorem 2. *The blockchain security game has an NE* $(x^*, y^*) = (0, j), \forall 1 \leq j \leq L$ *if*

$$G_0^j \geq \max_{1\leq i\leq L} \left(G_i^j - i\beta\right) \qquad (8)$$

$$C_0^j - j\beta \geq \max_{1\leq i\leq L} \left(C_0^i - i\beta, C_0^0 + R\right). \qquad (9)$$

Proof: The proof is similar to that of Theorem 1. ■

Theorem 3. *The blockchain security game has an NE* $(x^*, y^*) = (i, 0), \forall 1 \leq i \leq L$ *if*

$$G_i^0 - i\beta \geq \max_{1\leq j\leq L} \left(G_j^0 - j\beta, G_0^0 - R\right) \qquad (10)$$

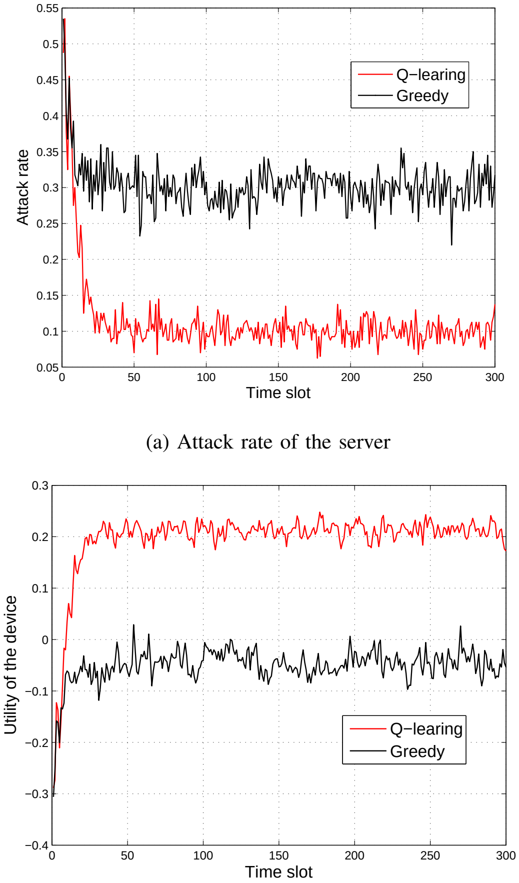
$$C_i^0 \geq \max_{1\leq j\leq L} \left(C_i^j - j\beta\right). \qquad (11)$$

Proof: The proof is similar to that of Theorem 1. ■

Theorem 4. *The blockchain security game has an NE*

Algorithm 1 Security scheme in the blockchain based network

```
Initialize $\alpha = 0.6, \delta = 0.8, s^{(0)}$, $\mathbf{Q} = \mathbf{0}$, $\mathbf{V} = \mathbf{0}$
for $k = 1, 2, ...$ do
    Choose $x^{(k)}$ via (16)
    if $x^{(k)} = i, 1 \leq i \leq L$ then
        Launch Level-$i$ attack against the server
    else
        Send a service request to the server
    end if
    Obtain utility $u_D^{(k)}$ and $y^{(k)}$
    Update $Q\left(s^{(k)}, x^{(k)}\right)$ via (14)
    Update $V\left(s^{(k)}\right)$ via (15)
    $s^{(k+1)} = y^{(k)}$
end for
```



(a) Attack rate of the server

(b) Utility of the device

Fig. 3: Performance of the dynamic security game, with $L = 1$, $R = 0.45$, $\beta = 0.6$, $\alpha = 0.6$, $\delta = 0.8$ and $\epsilon = 0.1$.

$(x^*, y^*) = (i, j), \forall 1 \leq i, j \leq L$ *if*

$$G_i^j - i\beta \geq \max_{1 \leq r \leq L} \left(G_r^j - r\beta, G_0^j\right) \tag{12}$$

$$C_i^j - j\beta \geq \max_{1 \leq r \leq L} \left(C_i^r - r\beta, C_i^0\right). \tag{13}$$

Proof: The proof is similar to that of Theorem 1. ■

Numerical results illustrating Theorems 1 through 4 are provided to illustrate the performance of the blockchain security game, with $R = 0.6$ and $L = 1$. As shown in Fig. 2, both the mobile device and the edge server tend to behave nicely as punishment factor increases, because the attack motivations of the device and the server are suppressed as the weight of the action record changes at 0.3 and 0.4, respectively. The turning points of the punishment factor $\beta = 0.3$ and 0.4 are given by Eqs. (5)-(6), (8)-(9) and (12)-(13). In this case, the utility of the device decreases by 66.3% as β increases from 0.1 to 0.3 because of the increasing punishment weight and security loss, and increases from -0.5 to 0.1 as β changes at 0.3. That is because the server chooses to follow the service request to avoid punishment.

V. Q-Learning Based Security Scheme

The repeated interactions between the mobile device and the edge server in the blockchain-based network can be formulated as a dynamic game. The decision making of the device has impacts on the future action policy of the edge server, and thus can be formulated as an MDP. Therefore, the device can apply reinforcement learning techniques such as Q-learning to derive its optimal strategy via trials without the knowledge of the current network condition.

In the dynamic game, the device decides to send a request or launch Level-i attack against the server based on the system state at time k denoted by $s^{(k)}$, which is chosen to be the network condition at the previous time, i.e., $s^{(k)} = y^{(k-1)}$. Let $Q(s, x)$ denote the Q-function of action x and state s. The value function $V(s)$ provides the maximum expected reward of the mobile device at system state s. The Q-function is updated according to Bellman iterative equation based on the immediate utility u_D as follows:

$$Q\left(s, x\right) \leftarrow (1 - \alpha) Q\left(s, x\right) + \alpha\left(u_D\left(s, x\right) + \delta V\left(s'\right)\right), \tag{14}$$

$$V\left(s\right) = \max_{x \in \mathbf{x}} Q\left(s, x\right), \tag{15}$$

where $\delta \in [0, 1]$ is the discount factor regarding the future reward, and $\alpha \in (0, 1]$ is the learning rate of the current experience.

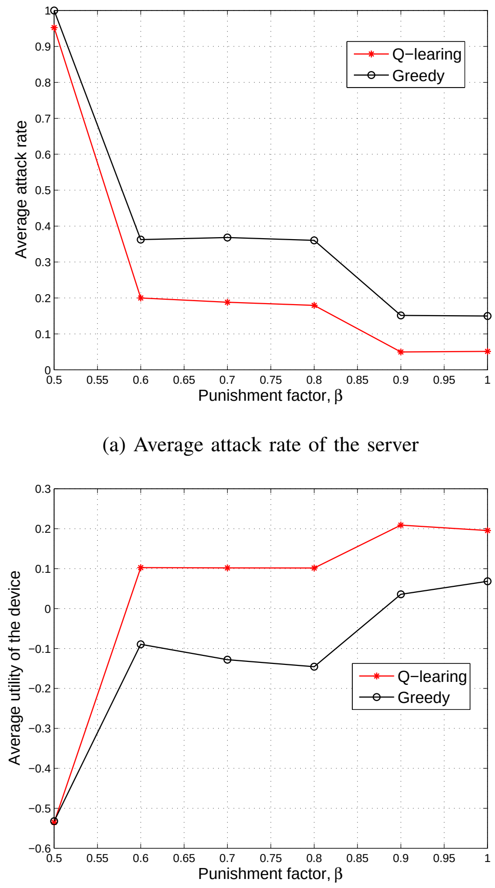
By applying the ϵ-greedy policy, the device chooses its action x to maximize its current Q-function as

$$\Pr(x = \tilde{x}) = \begin{cases} 1 - \epsilon, & \tilde{x} = \arg\max_{\mathbf{x}} Q\left(s, x\right) \\ \frac{\epsilon}{L}, & \text{o.w.} \end{cases} \tag{16}$$

The Q-learning based security scheme is summarized in Algorithm 1.

VI. Simulation Results

Simulations have been performed to evaluate the performance of the Q-learning based security scheme in the dynamic blockchain game. If not specified otherwise, we set $L = 1$,



(a) Average attack rate of the server

(b) Average utility of the device

Fig. 4: Performance of the dynamic security game averaged over 400 runs, with $L = 1$, $R = 0.45$, $\alpha = 0.6$, $\delta = 0.8$ and $\epsilon = 0.1$.

$R = 0.45$, $\beta = 0.6$, $\delta = 0.8$ and $\epsilon = 0.1$, and used a greedy strategy as benchmark, in which the action policy of the device is chosen to maximize the estimated immediate utility based on the previous network condition. The edge server chose strategy to maximize its utility according to the action history. The attack rate of the server denoted by R is the probability that the server launches an attack against the mobile device.

The security performance in the dynamic game is shown in Fig. 3, with $\beta = 0.6$. The attack rate of the server decreases over time, from 50% at the beginning of the game to 10% after 100 time slots, which is about 66.7% lower than that of the greedy strategy. Consequently, the utility of the mobile device increases over time from -0.3 at the beginning to 0.2 after 100 time slots. The reason is that the device learns the previous network condition and adjusts its policy via trials.

As shown in Fig. 4, the attack rate of the server decreases with the punishment factor, e.g., R increases by 94.7% if β changes from 0.5 to 1, because the attack motivation of the server is suppressed to avoid punishment. The attack rate is 5% with $\beta = 0.9$, which is 66.7% lower than the benchmark strategy. Consequently, the utility of the mobile device increases from -0.5 to 0.2, if β changes from 0.5 to 1.

VII. Conclusion

In this work, we have formulated a blockchain security game to investigate a punishment mechanism based on the action record on the blockchain. The NEs of the security game and the conditions under which the equilibria exist have been provided, showing that the mobile device and the edge server tend to behave nicely if the punishment weight is large. A Q-learning based security scheme has been proposed for the dynamic game to improve the network security performance, e.g., the attack rate of the server decreases by 66.7% compared with the benchmark strategy.

References

[1] W. Zhang, Y. Hu, Y. Zhang, and D. Raychaudhuri, “SEGUE: Quality of service aware edge cloud service migration,” in *Proc. IEEE Conf. Cloud Computing Technology and Science*, pp. 344–351, Luxembourg, Dec. 2016.

[2] M. Swan, “Blockchain: Blueprint for a new economy,” *O'Reilly Media, Inc.*, Feb. 2015.

[3] G. Zyskind and O. Nathan, “Decentralizing privacy: Using blockchain to protect personal data,” in *Proc. IEEE Security and Privacy (SP)*, pp. 180–184, San Jose, CA, May 2015.

[4] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, “Hawk: The blockchain model of cryptography and privacy-preserving smart contracts,” *Proc. IEEE Symposium Security and Privacy*, pp. 839–858, San Jose, CA, May 2015.

[5] M. Samaniego and R. Deters, “Hosting virtual iot resources on edge-hosts with blockchain,” in *Proc. IEEE Int. Conf. Computer and Information Technology (CIT)*, pp. 180–184, Nadi, Fiji, Dec. 2016.

[6] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, “Blockchain for iot security and privacy: The case study of a smart home,” *Pro. IEEE Int. Conf. Pervasive Computing and Communications Workshops*, pp. 618–623, Kona, Big Island, HI, Mar. 2017.

[7] L. Xiao, D. Xu, C. Xie, N. B. Mandayam, and H. V. Poor, “Cloud storage defense against advanced persistent threats: A prospect theoretic study,” *IEEE J. Selected Areas Commun.*, vol. 35, no. 3, Mar. 2017.

[8] A. Garnaev, M. Baykal-Gursoy, and H. V. Poor, “Security games with unknown adversarial strategies,” *IEEE Trans. Cybernetics*, vol. 46.

[9] 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.

[10] 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.

[11] Y. Yuan and F. Wang, “Towards blockchain-based intelligent transportation systems,” in *Proc. IEEE Int. Conf. Intelligent Transportation Systems (ITSC)*, pp. 2663–2668, Rio de Janeiro, Brazil, Dec. 2016.

[12] Y. Han, T. Alpcan, J. Chan, C. Leckie, and B. I. Rubinstein, “A game theoretical approach to defend against co-resident attacks in cloud computing: Preventing co-residence using semi-supervised learning,” *IEEE Trans. Information Forensics and Security*, vol. 11, no. 3, pp. 556–570, Mar. 2016.

[13] J. Chen, Q. Yuan, G. Xue, and R. Du, “Game-theory-based batch identification of invalid signatures in wireless mobile networks,” in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, pp. 262–270, HK, May 2015.

[14] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, “A mobile offloading game against smart attacks,” *IEEE Access*, vol. 4, pp. 2281–2291, May 2016.

[15] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra, “Dynamic defense strategy against advanced persistent threat with insiders,” in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, pp. 747–755, HK, May 2015.

[16] S. Sengupta, S. G. Wang, et al., “A game theoretic approach to strategy generation for moving target defense in web applications,” in *Proc. ACM Conf. Autonomous Agents and Multiagent Systems*, pp. 178–186, São Paulo, Brazil, May 2017.

[17] H. Zhang, P. Cheng, L. Shi, and J. Chen, “Optimal denial-of-service attack scheduling with energy constraint,” *IEEE Trans. Automatic Control*, vol. 60, no. 11, pp. 3023 – 3028, Nov. 2015.