Journal of the Franklin Institute 361 (2024) 164–176

Contents lists available at ScienceDirect

Journal of the Franklin Institute

journal homepage: www.elsevier.com/locate/fi

Security consensus control for multi-agent systems under DoS attacks via reinforcement learning method

Jinliang Liu [a,*], Yanhui Dong [b], Zhou Gu [c], Xiangpeng Xie [d], Engang Tian [e]

[a] School of Computer Science, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, 210044, China
[b] College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, Jiangsu, 210023, China
[c] College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, 210037, Jiangsu, China
[d] Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, 210023, China
[e] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, China

ARTICLE INFO

Keywords:
Multiagent systems (MASs)
Reinforcement learning (RL)
Denial-of-service (DoS) attacks
H_∞ consensus control

ABSTRACT

This paper is concerned with the security consensus control issue for discrete-time multiagent systems (MASs) on the basis of a reinforcement learning (RL) approach. Considering the effects of denial-of-service (DoS) attacks, a novel control protocol is proposed to deal with the H_∞ consensus problem. Firstly, a Q-learning algorithm is put forward under the directed graph, which can obtain the target gain matrices without any system dynamics information. In addition, the obtained gain matrices and Lyapunov function are employed to demonstrate that the MASs can reach security consensus. Moreover, the proof of H_∞ consensus under undirected graphs is derived using the designed Q-learning algorithm. In the end, the simulation experiments are given to verify the correctness of the designed control strategy.

1. Introduction

Recently, multiagent systems (MASs) have been increasingly concerned by many researchers due to the continuous expansion of their application range [1–6], such as artificial intelligence, biological ecology, and communication control. Especially, consensus behavior has sparked many valuable discussions as the most basic behavior of MASs [7–11]. The purpose of consensus control scheme is to achieve the desired consistency among all intelligent agents. To our knowledge, many research issues on consensus have been conducted in the secure communication environment [12,13]. In practical, there are many unsafe factors in the process of agents communication, such as cyber attacks, packet loss, and network delay, which will reduce communication quality and even cause system fluctuations [14]. Therefore, the security consensus problem of MASs suffering from network attacks needs to be paid attention to ensure normal system communication.

In MASs, network attacks are divided into two situations [15,16]. The first is that when a malicious cyber-attacks assault an agent in the communication networked diagram, the agent will be deleted, the second case is the communication interruption caused by cyber-attacks. As a common type of cyber-attacks, denial-of-service (DoS) attacks frequently occur in engineering practice [17–20]. It should be noticed that the open communication network can be unpredictably blocked under DoS attacks, which lead to the phenomenon that the signal cannot be normally sent to the controller [21–23]. Therefore, it is requisite to explore a security control policy for MASs. At present, some achievements have been made to resist the impact of DoS attacks. For example, an input-based event-triggered control strategy was put forward for MASs against DoS attacks in [24]. Liu et al. [25] concentrated on a secure leader-following controller design for MASs with replay and DoS attacks. Considering the impact of DoS attacks and actuator

* Corresponding author.
E-mail address: liujinliang@vip.163.com (J. Liu).

https://doi.org/10.1016/j.jfranklin.2023.11.032
Received 28 June 2023; Received in revised form 15 October 2023; Accepted 18 November 2023
Available online 25 November 2023

failures, the authors in [26] investigated a control strategy for nonlinear MASs by utilizing interval Takagi–Sugeno fuzzy model. Li et al. [27] and Feng et al. [28] discussed the security synchronization problem of discrete-time MASs under DoS attacks.

It is worth mentioning that the specific dynamics (A, B, D) needs to be obtained in the aforementioned results. However, the accurate system information is usually hard to acquire in the practical implementations. Thus, the aforesaid control schemes are inapplicable for the system with unknown dynamics. In an effort to solve this difficulty, with the assistance of reinforcement learning (RL) method [29–31], researchers have proposed several model-free algorithms to achieve the expected system stability of unknown system dynamics or optimal consensus of MASs [32–35]. The authors in [32] designed a consensus controller for MASs by using RL method. In [33], Long et al. put forward two Q-learning algorithms for discrete-time MASs to attain state feedback control. And the consensus issue was investigated for nonlinear MASs with external disturbance in [34]. Based on the Q-learning, the authors in [35] designed an optimal controller for unknown MASs. Although some effective control methods for the consensus of MASs have been presented in the above literature, these results did not take the network security into account. Therefore, the consistency of MASs subject to DoS attacks will be explored in virtue of a RL algorithm, which is the prominent innovation of this article.

Illuminated by the aforesaid investigation, a RL-based security consensus control policy is proposed for MASs subject to DoS attacks. The significant features of this article are outlined as follows:

(1) The published literature [30] have addressed the consensus control issue for discrete-time MASs. However, the adverse impact of cyber-attacks has not been taken into account. It is widely noticed that DoS attacks may degrade the system performance due to its attack manner. Towards this end, we endeavor to develop a secure consensus control scheme for MASs against DoS attacks.
(2) A model-free Q-learning algorithm is designed to derive the optimal control gain matrices. In contrast to [27], the target gain matrices can be iteratively acquired without any system dynamics information. With the assistance of the proposed algorithm, the desired H_∞ consensus for MASs can attain under the negative effect of DoS attacks.

The rest of this work is arranged as follows. In Section 2, considering DoS attacks, a new control protocol is constructed. In Section 3, the Q-learning method and Lyapunov function are used to derive the optimal controller. In Section 4, simulation results are given to demonstrate the effectiveness of the proposed approach. Finally, the conclusion is given in Section 5.

2. Problem formulation

2.1. Graph theory knowledges

Consider a graph $G = (\mathbb{V}, \mathbb{E}, \mathbb{D})$ with n agents. Define $\mathbb{V} = \{v_1, v_2, \dots, v_n\}$ as the node set, $\mathbb{E} = \{(v_i, v_j) : v_i, v_j \in \mathbb{V}\}$ is edge set, the row stochastic matrix is denoted as $\mathbb{D} = [d_{ij}] \in \mathcal{R}^{n\times n}$, which represents communication between all agents with $d_{ii} > 0$, $\sum_{j=1}^{n} d_{ij} = 1$ and

$$\begin{cases} d_{ij} > 0, if (v_i, v_j) \in \mathbb{E} \\ d_{ij} = 0, if (v_i, v_j) \neq \mathbb{E} \end{cases}$$

I_n represents the identity matrix. $I_n - \mathbb{D}$ is a particular Laplacian matrix, with

$$Re(\lambda_1(I_n - \mathbb{D})) < Re(\lambda_2(I_n - \mathbb{D})) < \cdots < Re(\lambda_n(I_n - \mathbb{D})).$$

2.2. System descriptions

Consider the following MASs with n agents

$$x_i(k+1) = Ax_i(k) + Bu_i(k) + E\omega_i(k) \tag{1}$$

where $i = 1, 2, \dots, n$; $x_i(k) \in \mathcal{R}^p$ represents the system state, $u_i(k) \in \mathcal{R}^q$ is the control input, $\omega_i(k) \in \mathcal{R}^s$ indicates the external disturbance of the ith agent, respectively. A, B and E are unknown system matrices with suitable dimensions.

Before proceeding to the main discussions of this article, the assumptions and lemmas are listed as

Assumption 1 (*[34]*). The (A, B) is stabilizable, (A, C) is observable, and $|\lambda_i(A)| \le 1 (i = 1, \dots, n)$.

Assumption 2 (*[36]*). G is a strongly connected and balanced directed graph or G is a connected undirected graph.

Assumption 3 (*[37]*). The transmission signal will be completely lost if the communication channel is under DoS attacks.

Lemma 1 (*[30]*). *For the directed graph G under Assumption 2, $\frac{2}{n(n-1)} \le Re(\lambda_2(I_n - \mathbb{D}))$ holds.*

Lemma 2 (*[30]*). *For the undirected graph G under Assumption 2, $\frac{4}{n(n-1)} \le \lambda_2(I_n - \mathbb{D})$ holds.*

2.3. Problem formulation

In practical implementations, it is difficult to accurately obtain the system dynamics. To overcome this difficulty, a Q-learning algorithm will be adopted to derive the gain matrices without knowing the system dynamics. In order to obtain the gain matrices, the following control protocol is proposed

$$u_i(k) = \alpha(k)K \sum_{j=1}^{n} d_{ij}(x_i(k) - x_j(k)) \tag{2}$$

where $K \in \mathcal{R}^{q\times p}$ is the controller gain, $\alpha(k)$ represents whether the DoS attacks are in presence at instant k and the specific meaning of $\alpha(k)$ is as follows:

$$\alpha(k) = \begin{cases} 0, & \text{if DoS attacks are active,} \\ 1, & \text{otherwise.} \end{cases} \tag{3}$$

The stochastic variable $\alpha(k)$ obeys the Bernoulli distribution taking values on $\{0, 1\}$ and the corresponding probabilities are

$$\begin{cases} Pr\{\alpha(k) = 1\} = \bar{\alpha}, \\ Pr\{\alpha(k) = 0\} = 1 - \bar{\alpha}, \end{cases}$$

where $\bar{\alpha} \in (0, 1)$ is a known constant. Apparently, $E\{\alpha(k)\} = E\{\alpha^2(k)\} = \bar{\alpha}$.

Remark 1. The control protocol shown in (2) is resultant from the influence of the DoS attacks. Specifically, $\alpha(k) = 1$ means the actuator successfully receives information from the controller and $\alpha(k) = 0$ otherwise.

Remark 2. Motivated by [38,39], the randomly occurring DoS attacks are modeled by a Bernoulli stochastic variable $\alpha(k)$. As stated in [40,41], some attack detection methods can be utilized to obtain the relevant information of DoS attacks through monitoring the communication network. Thus, the given probability $\bar{\alpha}$ can be acquired accordingly.

In the work, under the impact of DoS attacks, the optimal security consensus control problem is addressed by applying RL approach. Here, the design goal of this paper is presented as follows:

1. For all $x_i(0)$ and $\omega_i(k) = 0$, $\lim_{k\to+\infty} \|x_i(k) - x_j(k)\| = 0$.

2. For $x_i(0) = 0$, the following condition is satisfied:

$$\mathbf{E}\left\{\sum_{k=0}^{\infty}[x(k)^T Qx(k) + u(k)^T Ru(k)]\right\} \le \gamma^2 \mathbf{E}\left\{\sum_{k=0}^{\infty}\omega(k)^T\omega(k)\right\} \tag{4}$$

where $x(k) = [x_1^T(k), x_2^T(k), \dots, x_n^T(k)]^T$, $u(k) = [u_1^T(k), u_2^T(k), \dots, u_n^T(k)]^T$, $\omega(k) = [\omega_1^T(k), \omega_2^T(k), \dots, \omega_n^T(k)]^T$. Besides, $\gamma > 0$ denotes performance level and $Q \ge 0, R > 0$ are known weighting matrices.

Define a virtual control input $\varphi_i(k) = \alpha(k)\varphi_i^f(k)$ and a disturbance $f_i(k) = Lx_i(k)$, where $\varphi_i^f(k) = cKx_i(k)$ represents the auxiliary control variable, c is a constant to be determined, and L denotes the gain matrix to be devised. Then, the system (1) is expressed as

$$x_i(k+1) = Ax_i(k) + B\varphi_i(k) + Ef_i(k). \tag{5}$$

For simplicity, $x_i(k)$, $\varphi_i(k)$, $\varphi_i^f(k)$ and $f_i(k)$ are denoted as x_{ik}, φ_{ik}, φ_{ik}^f and f_{ik} in the following, respectively. The goal of the work is transformed into obtaining optimal φ_{ik}^* and the worst f_{ik}^*.

To achieve the goal, the following value function $V(x_{ik})$ is defined:

$$V(x_{ik}) = \mathbf{E}\left\{\sum_{i=k}^{\infty} J(x_{ik}, \varphi_{ik}, f_{ik})\right\} \tag{6}$$

where $J(\cdot)$ is the performance function with the following form:

$$J(x_{ik}, \varphi_{ik}, f_{ik}) = x_{ik}^T Qx_{ik} + \varphi_{ik}^T R\varphi_{ik} - \gamma^2 f_{ik}^T f_{ik}. \tag{7}$$

According to [42,43], regarding the virtual control input φ_{ik} and the disturbance f_{ik} as two players, the H_∞ consensus control problem in this article can be seen as a zero-sum game problem. According to the Bellman optimality principle, we are committed to solving a minmax problem under DoS attacks as

$$V^*(x_{ik}) = \min_{\varphi_{ik}} \max_{f_{ik}} \mathbf{E}\left\{J(x_{ik}, \varphi_{ik}, f_{ik})\right\} + V(x_{i(k+1)}). \tag{8}$$

Referring to the method in [44], the value function (6) has a quadratic form depending on x_{ik} as

$$V(x_{ik}) = \mathbf{E}\left\{x_{ik}^T Px_{ik}\right\} \tag{9}$$

where $P \ge 0$ is a symmetric matrix which will be designed later.

Then, the H_∞ consensus Q-function is given as

$$Q(x_{ik},\varphi_{ik},f_{ik}) = \mathbf{E}\left\{J(x_{ik},\varphi_{ik},f_{ik})\right\} + V(x_{i(k+1)}). \tag{10}$$

For the simplicity of the formulas, $Q(x_{ik},\varphi_{ik},f_{ik})$ in the following is represented as $\mathbb{Q}$. Define the augmented vector $\Xi_k = \begin{bmatrix} x_{ik}^T & (\varphi_{ik}^f)^T & f_{ik}^T \end{bmatrix}^T$. Combining the conditions (7) and (9), the expression (10) can be derived as

$$\begin{aligned}
\mathbb{Q} &= \mathbf{E}\left\{x_{ik}^T Q x_{ik} + \varphi_{ik}^T R \varphi_{ik} - \gamma^2 f_{ik}^T f_{ik}\right\} + \mathbf{E}\left\{x_{i(k+1)}^T P x_{i(k+1)}\right\} \\
&= \mathbf{E}\left\{x_{ik}^T Q x_{ik} + \varphi_{ik}^T R \varphi_{ik} - \gamma^2 f_{ik}^T f_{ik}\right\} \\
&\quad + \mathbf{E}\left\{(Ax_{ik} + B\varphi_{ik} + Ef_{ik})^T P (Ax_{ik} + B\varphi_{ik} + Ef_{ik})\right\} \\
&= \begin{bmatrix} x_{ik} \\ \varphi_{ik}^f \\ f_{ik} \end{bmatrix}^T \underbrace{\begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & N_{33} \end{bmatrix}}_{N} \begin{bmatrix} x_{ik} \\ \varphi_{ik}^f \\ f_{ik} \end{bmatrix}
\end{aligned} \tag{11}$$

where $N_{11} = A^T PA + Q$, $N_{12} = \bar{\alpha} A^T PB$, $N_{13} = A^T PE$, $N_{21} = \bar{\alpha} B^T PA$, $N_{22} = \bar{\alpha}(R + B^T PB)$, $N_{23} = \bar{\alpha} B^T PE$, $N_{31} = E^T PA$, $N_{32} = \bar{\alpha} E^T PB$, and $N_{33} = E^T PE - \gamma^2 I$.

Since the Q-function has a bearing on φ_{ik} and f_{ik}, target gain matrices K^* and L^* are solved by $\frac{\partial \mathbb{Q}}{\partial \varphi_{ik}} = 0$, $\frac{\partial \mathbb{Q}}{\partial f_{ik}} = 0$. By utilizing the formula (11), we have

$$\begin{aligned}
K^* &= m(N_{22} - N_{23}(N_{33})^{-1}N_{32})^{-1}(N_{23}(N_{33})^{-1}N_{31} - N_{21}) \\
L^* &= m(N_{33} - N_{32}(N_{22})^{-1}N_{23})^{-1}(N_{32}(N_{22})^{-1}N_{21} - N_{31})
\end{aligned} \tag{12}$$

with $m = \frac{1}{4n(n-1)}$.

Based on the expression of Ξ_k, we have

$$\mathbb{Q} = \Xi_k^T N \Xi_k. \tag{13}$$

Then, formula (13) is linearly parameterized as

$$\mathbb{Q}(\Xi_k) = \bar{N}^T \bar{\Xi}_{ik} \tag{14}$$

where

$$\bar{N}^T = [n_{11}, 2n_{12}, \ldots, 2n_{1l}, n_{22}, 2n_{23}, \ldots, 2n_{2l}, \ldots, n_{ll}]^T \tag{15}$$

and

$$\begin{aligned}
\bar{\Xi}_k = [&\Xi_{ik(1)}^2, \Xi_{ik(1)}\Xi_{ik(2)}, \ldots, \Xi_{ik(1)}\Xi_{ik(l)}, \\
&\Xi_{ik(2)}^2, \Xi_{ik(2)}\Xi_{ik(3)}, \ldots, \Xi_{ik(2)}\Xi_{ik(l)}, \ldots, \Xi_{ik(l)}^2]^T
\end{aligned} \tag{16}$$

in which n_{ij} is the element in the ith row and the jth column of matrix N, $i,j = 1,\ldots,l$, $l = p+q+s$, $\Xi_{ik(v)}$ is the vth component of vector Ξ_{ik}.

Then, the formula (14) is presented as

$$\bar{N}^T \bar{\Xi}_{ik} = x_{ik}^T Q x_{ik} + \bar{\alpha}(\varphi_{ik}^f)^T R \varphi_{ik}^f - \gamma^2 f_{ik}^T f_{ik} + \bar{N}^T \bar{\Xi}_{i(k+1)}. \tag{17}$$

According to the formula (17), Algorithm 1 will be put forward to derive matrix N online and obtain the optimal consensus controller.

Remark 3. Note that many existing available results about consensus control problems drew support from Q-learning algorithm for different systems on the premise of reliable communication channel, which is unrealistic in some cases. In this article, we aim to design a security consensus control method using the Q-learning algorithm for discrete-time MASs under DoS attacks, which is still challenging nowadays.

Remark 4. In Algorithm 1, probing noises p_{ik} and q_{ik} introduced in control input and external disturbance are inspired by the recent work [45], which can assure the condition of policy evaluation. Since the probing noises have not any impact on the formulated Q-function, the choice of probing noises is not a key issue. It should be noted that the sinusoidal function and exponential attenuation function are often used as the probing noises in many literatures [46]. Hence, the similar probing noises are also adopted in this paper.

Algorithm 1 Model-Free Q-Learning Algorithm

procedure System Initialization:
Set the iteration number $j = 0$, maximum iterations j_m.
Start with $N^0 > 0$, $K^0 = 0$, $L^0 = 0$, $\varphi_{ik}^f = cK^0 x_{ik} + p_{ik}$ and $f_{ik} = L^0 x_{ik} + q_{ik}$.

procedure Repeat:
1. Record $G \geq \frac{q(q+1)}{2}$ groups data of $(x_{ik}, \varphi_{ik}, f_{ik}, x_{i(k+1)}, \varphi_{i(k+1)}, f_{i(k+1)})$ at time k to form the data matrices $M \in \mathcal{R}^{\frac{l(l+1)}{2}\times G}$, $O \in \mathcal{R}^{G\times 1}$

$$\begin{cases} M = [\bar{\Xi}_{ik}^1, \bar{\Xi}_{ik}^2, \ldots, \bar{\Xi}_{ik}^G], \\ O = [J^1 + (\bar{N}^{j-1})^T \bar{\Xi}_{i(k+1)}^1, J^1 + (\bar{N}^{j-1})^T \bar{\Xi}_{i(k+1)}^2, \ldots, \\ \qquad J^G + (\bar{N}^{j-1})^T \bar{\Xi}_{i(k+1)}^G]^T. \end{cases} \tag{18}$$

2. Obtain N^j by

$$(\bar{N}^j)^T \bar{\Xi}_{ik} = x_{ik}^T Q x_{ik} + \bar{\alpha}(\varphi_{ik}^f)^T R \varphi_{ik}^f - \gamma^2 f_{ik}^T f_{ik} + (\bar{N}^{j-1})^T \bar{\Xi}_{i(k+1)}. \tag{19}$$

3. Update $\varphi_{ik}^f = cK^j x_{ik} + p_{ik}$ and $f_{ik} = L^j x_{ik} + q_{ik}$ using

$$\begin{aligned} K^j =& m(N_{22}^j - N_{23}^j (N_{33}^j)^{-1} N_{32}^j)^{-1} (N_{23}^j (N_{33}^j)^{-1} N_{31}^j - N_{21}^j), \\ L^j =& m(N_{33}^j - N_{32}^j (N_{22}^j)^{-1} N_{23}^j)^{-1} (N_{32}^j (N_{22}^j)^{-1} N_{21}^j - N_{31}^j). \end{aligned} \tag{20}$$

4. Stop
if $j > j_m$ **then**
Output the N^j, gain matrices K and L.
else
set $j = j + 1$ and go to step 1.

3. Main results

In what follows, the secure consensus of the concerned MASs can be ensured to achieve by virtue of the selected Lyapunov function from directed and undirected graph.

Theorem 1. *Under Assumptions 1–2, the MASs (5) under directed graph with formula (2) is able to reach secure consensus, where the optional control gain K^* as well as worst disturbance gain L^* are acquired from Algorithm 1 with*

$$4n(n-1) + 2\sqrt{4n^2(n-1)^2 - 3} \leq c < 8n(n-1). \tag{21}$$

Proof. Substituting the virtual control input $\varphi_{ik} = c\alpha(k)Kx_{ik}$ and the disturbance $f_{ik} = Lx_{ik}$ into the condition (11), we get

$$\begin{aligned} \mathbb{Q} =& \mathbf{E}\left\{x_{ik}^T Q x_{ik} + \varphi_{ik}^T R \varphi_{ik} - \gamma^2 f_{ik}^T f_{ik} + x_{i(k+1)}^T P x_{i(k+1)}\right\} \\ =& \mathbf{E}\left\{x_{ik}^T Q x_{ik} + \varphi_{ik}^T R \varphi_{ik} - \gamma^2 f_{ik}^T f_{ik}\right\} + \mathbf{E}\{(Ax_{ik} + B\varphi_{ik} + Ef_{ik})^T P(Ax_{ik} \\ & + B\varphi_{ik} + Ef_{ik})\} \\ =& \mathbf{E}\{x_{ik}^T Q x_{ik} + c^2\alpha^2(k) x_{ik}^T K^T R K x_{ik} - \gamma^2 x_{ik}^T L^T L x_{ik}\} + \mathbf{E}\{(Ax_{ik} + c\alpha(k) \\ & \times BKx_{ik} + ELx_{ik})^T P(Ax_{ik} + c\alpha(k)BKx_{ik} + ELx_{ik})\}. \end{aligned} \tag{22}$$

By calculation, the formula (22) is written as

$$\begin{aligned} \mathbb{Q} =& x_{ik}^T[Q + c^2\bar{\alpha}K^T RK - \gamma^2 L^T L + \mathcal{A}_{11} + c\bar{\alpha}\mathcal{A}_{12}K + \mathcal{A}_{13}L + c\bar{\alpha}K^T\mathcal{B}_{21} \\ & + c^2\bar{\alpha}K^T\mathcal{B}_{22}K + c\bar{\alpha}K^T\mathcal{B}_{23}L + L^T\mathcal{E}_{31} + c\bar{\alpha}L^T\mathcal{E}_{32}K + L^T\mathcal{E}_{33}L]x_{ik} \\ =& x_{ik}^T\mathscr{B}\left\{\begin{bmatrix} Q & & \\ & c^2\bar{\alpha}R & \\ & & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix}\begin{bmatrix} I & c\bar{\alpha} & I \\ c\bar{\alpha} & c^2\bar{\alpha} & c\bar{\alpha} \\ I & c\bar{\alpha} & I \end{bmatrix} P \right. \\ & \left. \times \begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix}\right\}\mathscr{B}^T x_{ik} \\ =& x_{ik}^T P x_{ik} \end{aligned} \tag{23}$$

with $\mathscr{B} = \begin{bmatrix} I & K^T & L^T \end{bmatrix}$, $\mathcal{A}_{11} = A^T PA$, $\mathcal{A}_{12} = A^T PB$, $\mathcal{A}_{13} = A^T PE$, $\mathcal{B}_{21} = B^T PA$, $\mathcal{B}_{22} = B^T PB$, $\mathcal{B}_{23} = B^T PE$, $\mathcal{E}_{31} = E^T PA$, $\mathcal{E}_{32} = E^T PB$, $\mathcal{E}_{33} = E^T PE$.

Then, we can get

$$x_{ik}^T \mathscr{B} \left\{ \begin{bmatrix} Q-P & & \\ & c^2\bar{\alpha}R & \\ & & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix} \begin{bmatrix} I & c\bar{\alpha} & I \\ c\bar{\alpha} & c^2\bar{\alpha} & c\bar{\alpha} \\ I & c\bar{\alpha} & I \end{bmatrix} P \right. \\ \left. \times \begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix} \right\} \mathscr{B}^T x_{ik} = 0. \tag{24}$$

In what follows, the discrete-time MASs will be proven to reach secure consensus with formula (2), where K is calculated through Algorithm 1. First, design an error function $z(k) = (\Theta(I_n - \mathbb{D}) \otimes I_n)x(k)$, and $I_n - \mathbb{D} = \Theta^{-1}\mathbb{R}\Theta$, with $\mathbb{R} \in R^{n\times n}$ being an upper-triangular matrix with $\lambda_i(I_n - \mathbb{D})$ as the diagonal terms. Then, the consensus is reached if $z(k) = 0$, i.e., $x_1(k) = \cdots = x_n(k)$. And the error function $z(k)$ is represented as

$$\begin{aligned} z(k+1) =& (\Theta(I_n - \mathbb{D}) \otimes I_n)x(k+1) \\ =& (\Theta(I_n - \mathbb{D}) \otimes I_n)[I_n \otimes A + (I_n - \mathbb{D}) \otimes \alpha(k)BK \\ &+ I_n \otimes EL]x(k) \\ =& [I_n \otimes A + \mathbb{R} \otimes \alpha(k)BK + I_n \otimes EL]z(k). \end{aligned} \tag{25}$$

Design an auxiliary system

$$\vec{z}(k+1) = [I_n \otimes A + \vec{\mathbb{R}} \otimes \alpha(k)BK + I_n \otimes EL]\vec{z}(k) \tag{26}$$

with $\vec{\mathbb{R}} = diag\{\lambda_1(I_n - \mathbb{D}), \dots, \lambda_n(I_n - \mathbb{D})\}$, $\vec{z}(k) = z(k)$.

When $\omega_i(k) = 0$, according to the formula (12), we can derive that $K = -mN_{22}^{-1}N_{21} = -m(R + B^T PB)^{-1}B^T PA$. Subsequently, it is apparently observed from the condition (21) that $c^2m^2 - 2cm \geq -3/[4n^2(n-1)^2]$.

Then, construct a Lyapunov function as

$$V(k) = \mathbf{E}\left\{\vec{z}^H(k)(I_n \otimes P)\vec{z}(k)\right\}. \tag{27}$$

According to the method used in [32], the following condition can be deduced by applying Lemma 1:

$$\begin{aligned} \Delta V(k) =& V(k+1) - V(k) \\ \leq& n\vec{z}_i^H(k)[\mathcal{A}_{11} - P + (4m^2 - \frac{4m}{n(n-1)})\bar{\alpha}\Pi]\vec{z}_i(k) \\ =& n\vec{z}_i^H(k)[\mathcal{A}_{11} - P - \frac{3}{4n^2(n-1)^2}\bar{\alpha}\Pi]\vec{z}_i(k) \\ \leq& n\vec{z}_i^H(k)[\mathcal{A}_{11} - P + (c^2m^2 - 2cm)\bar{\alpha}\Pi]\vec{z}_i(k) \end{aligned} \tag{28}$$

where $\Pi = A^T PB(B^T PB + R)^{-1}B^T PA$.

On account of the Eqs. (22) and (24), it is obvious that

$$\mathcal{A}_{11} - P + (c^2m^2 - 2cm)\bar{\alpha}\Pi < 0 \tag{29}$$

which implies $\Delta V(k) < 0$. Then, we have $\lim_{k\to+\infty} \|x_i(k) - x_j(k)\| = 0$, and the consensus control goal for the MASs (5) is attained.

For convenience, we denote the eigenvalue of $I_n - \mathbb{D}$ as λ_i. When $\omega_i(k) \neq 0$ and $x_i(0) = 0$, in terms of condition (4), with $V(\infty) = \lim_{k\to+\infty} V(k) \geq 0$, we can derive that

$$\begin{aligned} \mathbb{J} =& \mathbf{E}\left\{ \sum_{k=0}^{\infty}[\vec{z}^T(k)Q\vec{z}(k) + \alpha(k)\vec{u}^T(k)R\vec{u}(k) - \gamma^2\vec{\omega}^T(k)\vec{\omega}(k) + \triangle V(k)] \right. \\ &\left. - [V(\infty) - V(0)] \right\} \\ <& \mathbf{E}\left\{ \sum_{k=0}^{\infty}[\vec{z}^T(k)Q\vec{z}(k) + \alpha(k)\vec{u}^T(k)R\vec{u}(k) - \gamma^2\vec{\omega}^T(k)\vec{\omega}(k) + \triangle V(k)] \right\} \\ =& \mathbf{E}\left\{ \sum_{k=0}^{\infty}[\vec{z}^T(k)Q\vec{z}(k) + \alpha(k)\vec{u}^T(k)R\vec{u}(k) - \gamma^2\vec{\omega}^T(k)\vec{\omega}(k) + \vec{z}^H(k)\mathscr{A}\vec{z}(k)] \right\} \\ =& \sum_{k=0}^{\infty}\sum_{i=1}^{n} \left\{ \vec{z}_i^T(k)Q\vec{z}_i(k) + \bar{\alpha}\vec{u}_i^T(k)R\vec{u}_i(k) - \gamma^2\vec{\omega}_i^T(k)\vec{\omega}_i(k) + \vec{z}_i^H(k)(\mathcal{A}_{11} \right. \\ &+ \lambda_i\bar{\alpha}\mathcal{A}_{12}K + \mathcal{A}_{13}L + \lambda_i\bar{\alpha}K^T\mathcal{B}_{21} + |\lambda_i|^2\bar{\alpha}K^T\mathcal{B}_{22}K + L^T\mathcal{E}_{31} + \lambda_i\bar{\alpha}K^T\mathcal{B}_{23}L \\ &\left. + \lambda_i\bar{\alpha}L^T\mathcal{E}_{32}K + L^T\mathcal{E}_{33}L - P)\vec{z}_i(k) \right\} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\infty}\sum_{i=1}^{n}\Bigg\{ \vec{z}_i^T(k)Q\vec{z}_i(k) + \bar{\alpha}\vec{u}_i^T(k)R\vec{u}_i(k) - \gamma^2\vec{\omega}_i^T(k)\vec{\omega}_i(k) \\
&+ \vec{z}_i^H(k)\mathscr{B}\Bigg\{\begin{bmatrix} -P & & \\ & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix}\begin{bmatrix} I & \lambda_i\bar{\alpha} & I \\ \lambda_i\bar{\alpha} & |\lambda_i|^2\bar{\alpha} & \lambda_i\bar{\alpha} \\ I & \lambda_i\bar{\alpha} & I \end{bmatrix} \\
&\times P\begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix}\Bigg\}\mathscr{B}^T\vec{z}_i(k)\Bigg\} \\
&= \sum_{k=0}^{\infty}\sum_{i=1}^{n}\Bigg\{ \vec{z}_i^H(k)\mathscr{B}\Bigg\{\begin{bmatrix} Q-P & & \\ & |\lambda_i|^2\bar{\alpha}R & \\ & & -\gamma^2 I \end{bmatrix} \\
&+ \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix}\begin{bmatrix} I & \lambda_i\bar{\alpha} & I \\ \lambda_i\bar{\alpha} & |\lambda_i|^2\bar{\alpha} & \lambda_i\bar{\alpha} \\ I & \lambda_i\bar{\alpha} & I \end{bmatrix} P\begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix}\Bigg\}\mathscr{B}^T\vec{z}_i(k)\Bigg\}
\end{aligned}
\tag{30}
$$

where $\mathscr{A} = [I_n \otimes A + \vec{\mathbb{R}} \otimes \alpha(k)BK + I_n \otimes EL]^T(I_n \otimes P)[I_n \otimes A + \vec{\mathbb{R}} \otimes \alpha(k)BK + I_n \otimes EL] - I_n \otimes P$, $\vec{u}(k) = (\vec{\mathbb{R}} \otimes K)\vec{z}(k)$, $\vec{\omega}(k) = L\vec{z}(k)$.

From the formula (4), one can obtain

$$
\mathbf{E}\left\{\sum_{k=0}^{\infty}[\vec{z}(k)^T Q\vec{z}(k) + \vec{u}(k)^T R\vec{u}(k)]\right\} \le \gamma^2\mathbf{E}\left\{\sum_{k=0}^{\infty}\vec{\omega}(k)^T\vec{\omega}(k)\right\}. \tag{31}
$$

Based on (24), one has

$$
\begin{aligned}
&\begin{bmatrix} 0 & & \\ & -c^2\bar{\alpha}R & \\ & & \gamma^2 I \end{bmatrix} + \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix}\begin{bmatrix} -I & -c\bar{\alpha} & -I \\ -c\bar{\alpha} & -c^2\bar{\alpha} & -c\bar{\alpha} \\ -I & -c\bar{\alpha} & -I \end{bmatrix} P\begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix} \\
&= \begin{bmatrix} Q-P & & \\ & 0 & \\ & & 0 \end{bmatrix}.
\end{aligned}
\tag{32}
$$

Substituting the formula (32) into (30), we get

$$
\begin{aligned}
\mathbb{J} &< \sum_{k=0}^{\infty}\sum_{i=1}^{n}\vec{z}_i^H(k)\mathscr{B}\Bigg\{\begin{bmatrix} 0 & & \\ & (|\lambda_i|^2 - c^2)\bar{\alpha}R & \\ & & 0 \end{bmatrix} + \begin{bmatrix} A^T & & \\ & B^T & \\ & & E^T \end{bmatrix} \\
&\times \begin{bmatrix} 0 & (\lambda_i - c)\bar{\alpha} & 0 \\ (\lambda_i - c)\bar{\alpha} & (|\lambda_i|^2 - c^2)\bar{\alpha} & (\lambda_i - c)\bar{\alpha} \\ 0 & (\lambda_i - c)\bar{\alpha} & 0 \end{bmatrix} P\begin{bmatrix} A & & \\ & B & \\ & & E \end{bmatrix}\Bigg\}\mathscr{B}^T\vec{z}_i(k).
\end{aligned}
\tag{33}
$$

Let

$$
S = \begin{bmatrix} 0 & (\lambda_i - c)\bar{\alpha} & 0 \\ (\lambda_i - c)\bar{\alpha} & (|\lambda_i|^2 - c^2)\bar{\alpha} & (\lambda_i - c)\bar{\alpha} \\ 0 & (\lambda_i - c)\bar{\alpha} & 0 \end{bmatrix}. \tag{34}
$$

Obviously, the matrix S is negative definite and $|\lambda_i|^2 - c^2 < 0$ with $|\lambda_i| < 1$, thus, $\mathbb{J} < 0$. Then, the MASs (5) can achieve consensus control, which completes the proof. ■

In what follows, we will discuss the case of the undirected graph. Using Algorithm 1 to compute the gain matrix K, the value of m needs to be changed. The relevant results and proof are presented as follows:

Theorem 2. *Under Assumptions 1–2, the MASs (5) under undirected graph with formula (2) can get secure consensus, where the optimal control gain K^* as well as worst disturbance gain L^* are acquired from Algorithm 1 with*

$$
4n(n-1) + 2\sqrt{4n^2(n-1)^2 - 7} \le c < 8n(n-1). \tag{35}
$$

Proof. Choose a Lyapunov function as

$$
V(k) = \mathbf{E}\left\{\vec{z}^H(k)(I_n \otimes P)\vec{z}(k)\right\}. \tag{36}
$$

Based on Eq. (12), we can derive that $K = -mN_{22}^{-1}N_{21} = -m(R + B^T PB)^{-1}B^T PA$ when $\omega_i(k) = 0$. With formula (32) and $m = \frac{1}{4n(n-1)}$, we can deduce $c^2m^2 - 2cm \ge -7/[4n^2(n-1)^2]$.

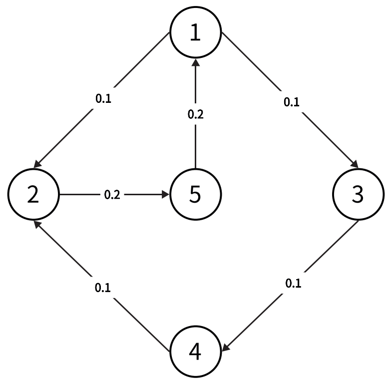

Fig. 1. Directed topology.

In light of the approach in [32,34], the following results can be obtained:

$$\begin{aligned}\Delta V(k) &= V(k+1) - V(k)\\ &\le \sum_{i=1}^{n} \vec{z}_i^H(k)[\mathcal{A}_{11} + \bar{\alpha} m^2 |\lambda_i|^2 \Pi - 2\bar{\alpha} m \lambda_i \Pi - P]\vec{z}_i(k).\end{aligned} \tag{37}$$

On the basis of Lemma 2 and the demonstration given in [32], it is evidently concluded that $\lambda_i \ge 4/[n(n-1)]$ and $|\lambda_i| \le 2$. Then, $\Delta V(k)$ is represented by

$$\begin{aligned}\Delta V(k) &\le n\vec{z}_i^H(k)[\mathcal{A}_{11} - P + (4m^2 - \frac{8m}{n(n-1)})\bar{\alpha}\Pi]\vec{z}_i(k)\\ &= n\vec{z}_i^H(k)[\mathcal{A}_{11} - P - \frac{7}{4n^2(n-1)^2}\bar{\alpha}\Pi]\vec{z}_i(k)\\ &\le n\vec{z}_i^H(k)[\mathcal{A}_{11} - P + (c^2m^2 - 2cm)\bar{\alpha}\Pi]\vec{z}_i(k).\end{aligned} \tag{38}$$

From (22) and (24), we have

$$\mathcal{A}_{11} - P + (c^2m^2 - 2cm)\bar{\alpha}\Pi < 0 \tag{39}$$

which implies $\Delta V(k) < 0$. Then, $\lim_{k\to+\infty} \|x_i(k) - x_j(k)\| = 0$ can be obtained, that is the MASs (5) can achieve consensus control.

On the other hand, for the situation that $\omega_i(k) \neq 0$, the relevant proof can also be finished similar to the proof of Theorem 1. Thus, the consensus control for MASs (5) under undirected graph can be achieved, which completes the proof. ■

4. Simulation examples

In the section, simulation results are shown to illustrate the validity of the designed secure consensus control method. In addition, the directed graph and the undirected graph are considered respectively.

Consider the MASs with five agents, the system matrices are

$$A = \begin{bmatrix} 0.95 & 0.1 \\ -0.8 & 0.3 \end{bmatrix}, B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, E = \begin{bmatrix} 0.15 \\ -0.4 \end{bmatrix}$$

which satisfies Assumption 1.

Case 1. As shown in Fig. 1, G is a directed graph, where

$$\mathbb{D} = \begin{bmatrix} 0.8 & 0 & 0 & 0 & 0.2 \\ 0.1 & 0.8 & 0 & 0.1 & 0 \\ 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0.2 & 0 & 0 & 0.8 \end{bmatrix}.$$

The original states are selected as $x_1(0) = [-3.1, -3.2]^T$, $x_2(0) = [-2.1, 2.2]^T$, $x_3(0) = [1.1, -4.3]^T$, $x_4(0) = [2.1, -2.5]^T$, $x_5(0) = [-1.0, -3.4]^T$. Set initial parameters as $\gamma = 0.95, m = \frac{1}{80}, c = 159.94$, $Q = 100, R = 10$, $K^0 = [0, 0], L^0 = [0, 0]$. The occurrence probability of DoS attacks are set as $1 - \bar{\alpha} = 0.2$.

By Algorithm 1, the gain matrices K and L are eventually computed as

$$K = \begin{bmatrix} 0.0207 & 0.0067 \end{bmatrix}, \ L = \begin{bmatrix} -0.0808 & -0.0084 \end{bmatrix}.$$

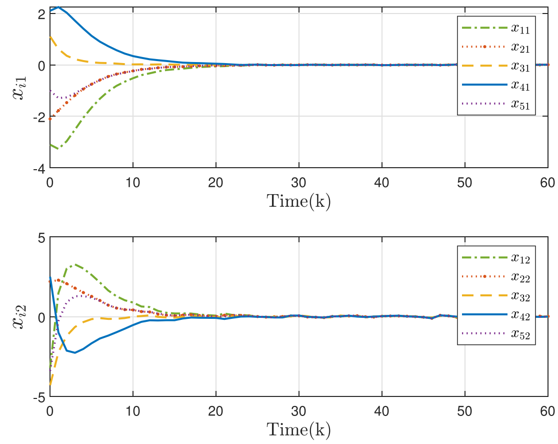

Fig. 2. The responses of x_i in Case 1.

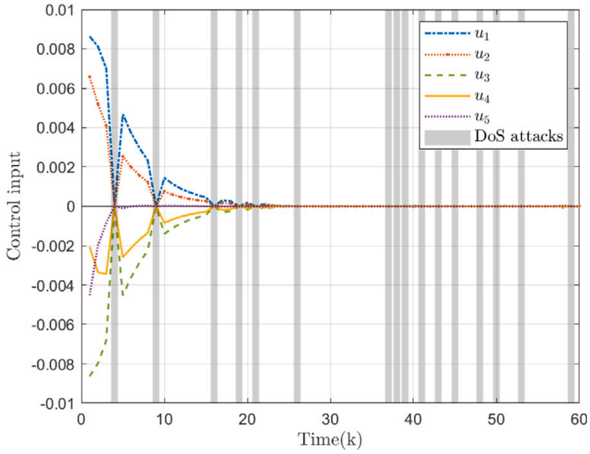

Fig. 3. Control input u_i in Case 1.

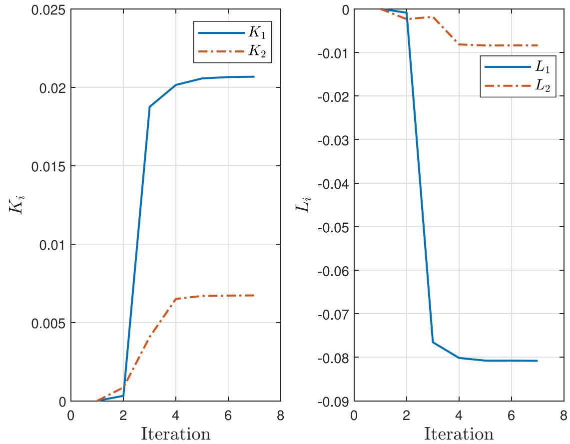

Fig. 4. Convergence of gain matrices K and L in Case 1.

Using the above control policy generated by Q-learning algorithm, Fig. 2 shows the responses of system states x_i in Case 1, which illustrates that the system can gradually reach consensus through our designed method. Fig. 3 describes the control input u_i in the

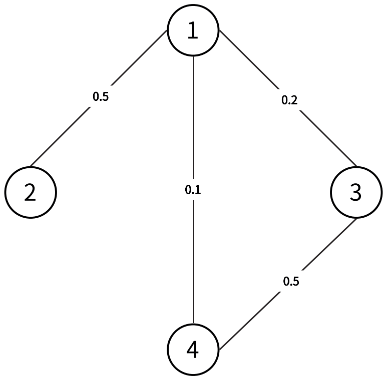

Fig. 5. Undirected topology.

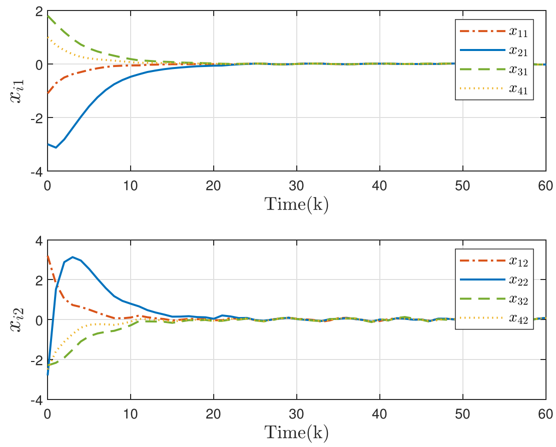

Fig. 6. The responses of x_i in Case 2.

presence of DoS attacks. Moreover, the convergence process of gain matrices K and L are depicted in Fig. 4. Based on the above observations, it is evident to conclude that the designed security consensus control strategy is effective under the directed graph.

Case 2. In this case, G is undirected as Fig. 5, where

$$\mathbb{D} = \begin{bmatrix} 0.2 & 0.5 & 0.2 & 0.1 \\ 0.5 & 0.5 & 0 & 0 \\ 0.2 & 0 & 0.3 & 0.5 \\ 0.1 & 0 & 0.5 & 0.4 \end{bmatrix}.$$

Then, the original states are selected as $x_1(0) = [-1.1, 3.2]^T$, $x_2(0) = [-3.0, -2.8]^T$, $x_3(0) = [1.8, -2.3]^T$, $x_4(0) = [1, -2.5]^T$. Other initial parameters are selected to be $\gamma = 0.95, m = \frac{1}{48}, c = 95.8$, $Q = 100$, $R = 10$, $K^0 = [0, 0]$, $L^0 = [0, 0]$. The occurrence probability of DoS attacks is $1 - \bar{\alpha} = 0.2$.

According to Algorithm 1, the matrices K and L are computed as follows:

$$K = \begin{bmatrix} 0.0345 & 0.0112 \end{bmatrix},\ L = \begin{bmatrix} -0.1347 & -0.0140 \end{bmatrix}.$$

In the following, we will present the case where the topology is an undirected graph. The system states x_i and the control input u_i are plotted in Figs. 6 and 7, respectively. It can be observed from Fig. 6 that consensus performance of the MASs can be satisfied gradually despite the DoS attacks. Fig. 8 represents the learning process of control gain K and the disturbance gain L, respectively. Apparently, when the malicious DoS attacks occur, the data transmission can be blocked such that the control input u_i becomes zero. Under such a negative impact, the proposed security consensus goal can still be achieved according to Fig. 6. Thus, the effectiveness of the applied control scheme is validated from the undirected graph.

5. Conclusion

In the article, the issue of security consensus control has been discussed for the MASs under DoS attacks using RL methods. Considering the DoS attacks, a new control protocol is proposed to solve the H_∞ consensus problem. Based on the topological

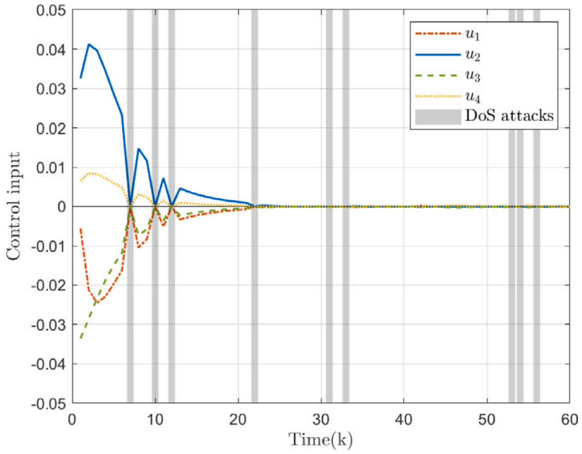

Fig. 7. Control input u_i in Case 2.

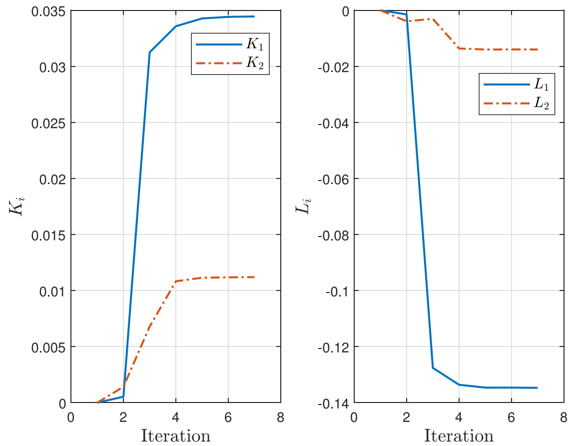

Fig. 8. Convergence of gain matrices K and L in Case 2.

structure of graphs, a Q-learning algorithm for the system has been put forward, which can obtain the optimal gain matrices without any system dynamics information. In the end, the simulation experiments have been given to demonstrate the correctness of the designed strategy. Further research directions will include the security controller design for MASs subject to multiple cyber attacks, which are consisted of DoS attacks, deception attacks and so on. Meanwhile, taking the restricted communication resource into account, the secure event-triggered control scheme will be investigated for MASs.

CRediT authorship contribution statement

Jinliang Liu: Conceptualization, Resources, Supervision, Writing – review & editing. **Yanhui Dong:** Software, Writing – original draft, Data curation. **Zhou Gu:** Methodology, Visualization. **Xiangpeng Xie:** Validation, Formal analysis, Funding acquisition. **Engang Tian:** Investigation, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants 61973152 and 62373252.

References

[1] Y. Hao, L. Liu, Event-triggered H_∞ output consensus of heterogeneous linear multi-agent systems, J. Franklin Inst. B 359 (16) (2022) 9056–9078.
[2] Y. Tang, D. Zhang, P. Shi, W. Zhang, F. Qian, Event-based formation control for nonlinear multiagent systems under DoS attacks, IEEE Trans. Automat. Control 66 (1) (2021) 452–459.
[3] H. Li, J. Cao, Event-triggered group consensus for one-sided Lipschitz multi-agent systems with input saturation, Commun. Nonlinear Sci. Numer. Simul. 121 (2023) 107234.
[4] H. Zhang, Z. Gao, Y. Wang, Y. Cai, Leader-following exponential consensus of fractional-order descriptor multiagent systems with distributed event-triggered strategy, IEEE Trans. Syst. Man Cybern.: Syst. 52 (6) (2022) 3967–3979.
[5] S. Chen, Z. Zou, Z. Zhang, L. Zhao, Fixed-time scaled consensus of multi-agent systems with input delay, J. Franklin Inst. B 360 (12) (2023) 8821–8840.
[6] D. Zhang, Y. Tang, Z. Ding, F. Qian, Event-based resilient formation control of multiagent systems, IEEE Trans. Cybern. 51 (5) (2021) 2490–2503.
[7] S. Fan, P. He, P. Shi, Bipartite consensus of multi-agent systems with matched uncertainty via fully distributed edge-based event-triggered mechanism, J. Franklin Inst. B 360 (12) (2023) 8585–8613.
[8] E. Arabi, D. Panagou, Adaptive control of second-order safety-critical multiagent systems with nonlinear dynamics, IEEE Trans. Control Netw. Syst. 9 (4) (2022) 1911–1922.
[9] Y. Tang, X. Xing, H.R. Karimi, L. Kocarev, J. Kurths, Tracking control of networked multi-agent systems under new characterizations of impulses and its applications in robotic systems, IEEE Trans. Ind. Electron. 63 (2) (2016) 1299–1307.
[10] L. Zha, R. Liao, J. Liu, X. Xie, J. Cao, L. Xiong, Finite-time adaptive event-triggered asynchronous state estimation for Markov jump systems with cyber-attacks, Internat. J. Robust Nonlinear Control 32 (2) (2022) 583–599.
[11] H. Zhu, J. Liu, Z. Zhang, S. Zhang, F. Qu, X. Liu, H_∞ consensus of multi-agent systems under hybrid cyber attacks via a sampled-data-based dynamic event-triggered resilient consensus protocol, J. Franklin Inst. B 360 (13) (2023) 9924–9949.
[12] G. Zhao, L. Cao, X. Li, Q. Zhou, Observer-based dynamic event-triggered control for nonstrict-feedback stochastic nonlinear multiagent systems, Appl. Math. Comput. 430 (2022) 127289.
[13] S. Yang, Y. Wen, B. Qiao, K. Wang, X. Su, Fault detection filter design for nonlinear singular systems with Markovian jump parameters, IEEE Syst. J. 15 (3) (2021) 4168–4176.
[14] C. Deng, C. Wen, Distributed resilient observer-based fault-tolerant control for heterogeneous multiagent systems under actuator faults and DoS attacks, IEEE Trans. Control Netw. Syst. 7 (3) (2020) 1308–1318.
[15] S. Du, Y. Wang, L. Dong, X. Li, Secure consensus of multiagent systems with DoS attacks via a graph-based approach, Inform. Sci. 570 (2021) 94–104.
[16] N. Zhao, P. Shi, W. Xing, R.K. Agarwal, Resilient event-triggered control for networked cascade control systems under denial-of-service attacks and actuator saturation, IEEE Syst. J. 16 (1) (2022) 1114–1122.
[17] G. Wu, G.-H. Yang, H. Wang, ISS control synthesis of T-S fuzzy systems with multiple transmission channels under denial of service, J. Franklin Inst. B 358 (6) (2021) 3010–3032.
[18] L. Zha, R. Liao, J. Liu, X. Xie, E. Tian, J. Cao, Dynamic event-triggered output feedback control for networked systems subject to multiple cyber attacks, IEEE Trans. Cybern. 52 (12) (2022) 13800–13808.
[19] J. Liu, Z.-G. Wu, D. Yue, J.H. Park, Stabilization of networked control systems with hybrid-driven mechanism and probabilistic cyber attacks, IEEE Trans. Syst. Man Cybern.: Syst. 51 (2) (2021) 943–953.
[20] J. Liu, Y. Wang, L. Zha, X. Xie, E. Tian, An event-triggered approach to security control for networked systems using hybrid attack model, Internat. J. Robust Nonlinear Control 31 (12) (2021) 5796–5812.
[21] S. Hu, X. Ge, Y. Li, X. Chen, X. Xie, D. Yue, Resilient load frequency control of multi-area power systems under DoS attacks, IEEE Trans. Inf. Forensics Secur. 18 (2023) 936–947.
[22] X. Zhao, S. Zou, Z. Ma, Decentralized resilient H_∞ load frequency control for cyber-physical power systems under DoS attacks, IEEE/CAA J. Autom. Sin. 8 (11) (2021) 1737–1751.
[23] X. Chen, S. Hu, Y. Li, D. Yue, C. Dou, L. Ding, Co-estimation of state and FDI attacks and attack compensation control for multi-area load frequency control systems under FDI and DoS attacks, IEEE Trans. Smart Grid 13 (3) (2022) 2357–2368.
[24] Y. Xu, M. Fang, Z.-G. Wu, Y.-J. Pan, M. Chadli, T. Huang, Input-based event-triggering consensus of multiagent systems under denial-of-service attacks, IEEE Trans. Syst. Man Cybern.: Syst. 50 (4) (2020) 1455–1464.
[25] J. Liu, T. Yin, D. Yue, H.R. Karimi, J. Cao, Event-based secure leader-following consensus control for multiagent systems with multiple cyber attacks, IEEE Trans. Cybern. 51 (1) (2021) 162–173.
[26] Z. Zhang, J. Dong, Fault-tolerant containment control for IT2 fuzzy networked multiagent systems against denial-of-service attacks and actuator faults, IEEE Trans. Syst. Man Cybern.: Syst. 52 (4) (2022) 2213–2224.
[27] X.-M. Li, D. Yao, P. Li, W. Meng, H. Li, R. Lu, Secure finite-horizon consensus control of multiagent systems against cyber attacks, IEEE Trans. Cybern. 52 (9) (2022) 9230–9239.
[28] S. Feng, H. Ishii, Dynamic quantized consensus of general linear multiagent systems under denial-of-service attacks, IEEE Trans. Control Netw. Syst. 9 (2) (2022) 562–574.
[29] W. Wang, X. Chen, H. Fu, M. Wu, Model-free distributed consensus control based on actor-critic framework for discrete-time nonlinear multiagent systems, IEEE Trans. Syst. Man Cybern.: Syst. 50 (11) (2020) 4123–4134.
[30] M. Long, H. Su, Z. Zeng, Output-feedback global consensus of discrete-time multiagent systems subject to input saturation via Q-learning method, IEEE Trans. Cybern. 52 (3) (2022) 1661–1670.
[31] B.C. Chung, D.-H. Cho, Semidynamic cell-clustering algorithm based on reinforcement learning in cooperative transmission system, IEEE Syst. J. 12 (4) (2018) 3853–3856.
[32] Y. Liu, H. Su, General second-order consensus of discrete-time multiagent systems via Q-learning method, IEEE Trans. Syst. Man Cybern.: Syst. 52 (3) (2022) 1417–1425.
[33] M. Long, H. Su, Z. Zeng, Model-free algorithms for containment control of saturated discrete-time multiagent systems via Q-learning method, IEEE Trans. Syst. Man Cybern.: Syst. 52 (2) (2022) 1308–1316.
[34] C. An, H. Su, S. Chen, H_∞ Consensus for discrete-time fractional-order multi-agent systems with disturbance via Q-learning in zero-sum games, IEEE Trans. Netw. Sci. Eng. 9 (4) (2022) 2803–2814.
[35] Z. Peng, R. Luo, J. Hu, K. Shi, S.K. Nguang, B.K. Ghosh, Optimal tracking control of nonlinear multiagent systems using internal reinforce Q-learning, IEEE Trans. Neural Netw. Learn. Syst. 33 (8) (2022) 4043–4055.
[36] C. Dou, D. Yue, X. Li, Y. Xue, MAS-based management and control strategies for integrated hybrid energy system, IEEE Trans. Ind. Inform. 12 (4) (2016) 1332–1349.
[37] E. Mousavinejad, X. Ge, Q.-L. Han, F. Yang, L. Vlacic, Resilient tracking control of networked control systems under cyber attacks, IEEE Trans. Cybern. 51 (4) (2021) 2107–2119.
[38] Y. Li, F. Song, J. Liu, X. Xie, E. Tian, Software defined event-triggering control for large-scale networked systems subject to stochastic cyber attacks, IEEE Trans. Control Netw. Syst. 10 (3) (2023) 1531–1541.

[39] Y. Tan, Y. Yuan, X. Xie, E. Tian, J. Liu, Observer-based event-triggered control for interval type-2 fuzzy networked system with network attacks, IEEE Trans. Fuzzy Syst. 31 (8) (2023) 2788–2798.
[40] D. Zhang, Q.-G. Wang, G. Feng, Y. Shi, A.V. Vasilakos, A survey on attack detection, estimation and control of industrial cyber-physical systems, ISA Trans. 116 (2021) 1–16.
[41] J. Liu, N. Zhang, Y. Li, X. Xie, H_∞ filter design for discrete-time networked systems with adaptive event-triggered mechanism and hybrid cyber attacks, J. Franklin Inst. B 358 (17) (2021) 9325–9345.
[42] B. Kiumarsi, F.L. Lewis, Z.-P. Jiang, H_∞ control of linear discrete-time systems: Off-policy reinforcement learning, Automatica 78 (2017) 144–152.
[43] C. Wu, X. Li, W. Pan, J. Liu, L. Wu, Zero-sum game-based optimal secure control under actuator attacks, IEEE Trans. Automat. Control 66 (8) (2021) 3773–3780.
[44] Y. Ren, Q. Wang, Z. Duan, Output-feedback Q-learning for discrete-time linear H_∞ tracking control: A stackelberg game approach, Internat. J. Robust Nonlinear Control 32 (12) (2022) 6805–6828.
[45] R. Zhang, K. Xiong, W. Guo, X. Yang, P. Fan, K.B. Letaief, Q-learning-based adaptive power control in wireless RF energy harvesting heterogeneous networks, IEEE Syst. J. 15 (2) (2021) 1861–1872.
[46] Y. Peng, Q. Chen, W. Sun, Reinforcement Q-learning algorithm for H_∞ tracking control of unknown discrete-time linear systems, IEEE Trans. Syst. Man Cybern.: Syst. 50 (11) (2020) 4109–4122.