Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Reinforcement learning based privacy-preserving consensus tracking control of nonstrict-feedback discrete-time multi-agent systems*#

Yang YANG†1, Fanming HUANG[1], Dong YUE†‡1,2

[1]*College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*

[2]*Advanced Technology Institute for Carbon Neutrality, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*

†E-mail: yyang@njupt.edu.cn; medongy@vip.163.com

Received Aug. 6, 2023; Revision accepted Nov. 22, 2023; Crosschecked Mar. 6, 2025

Abstract: This paper investigates a privacy-preserving consensus tracking problem for a class of nonstrict-feedback discrete-time multi-agent systems (MASs). An improved Liu cryptosystem is developed to alleviate the errors between encryption and decryption on the plaintext, which ensures satisfactory recovery of the plaintext information. A reinforcement learning (RL) technique is then employed to compensate for unknown dynamics and errors between true signals and decrypted ones. Based on the backstepping and graph theory, an RL-based privacy-preserving consensus tracking control strategy is further designed. By virtue of graph theory and Lyapunov stability theory, it is shown that the consensus tracking errors and all signals in the MAS are ultimately bounded. Finally, simulation examples are presented for verification of the effectiveness of the control strategy.

Key words: Multi-agent systems; Consensus tracking; Privacy-preserving; Reinforcement learning
https://doi.org/10.1631/FITEE.2300532 **CLC number:** TP13

1 Introduction

In recent years, consensus control of multi-agent systems (MASs) has attracted extensive attention in the control community (Zhang HG et al., 2017; Li HY et al., 2021; He et al., 2022; Ju et al., 2022; Wen and Li, 2022; Yang XD et al., 2022; Li JN et al., 2023; Zhang XM et al., 2023). This is primarily because consensus control finds its broad applications in diverse areas, as discussed by multiple autonomous vehicles (Shahvali et al., 2018; Peng et al., 2021; Ge XH et al., 2022, 2023, 2024; Xie et al., 2022), manipulators (Yu et al., 2019), aerospace engineering (Sakthivel et al., 2019), mobile robots (Ding et al., 2020; Ning et al., 2023), and power systems (Li P et al., 2022). Although MASs provide flexible and convenient platforms for industry and military applications, they are vulnerable to sensitive information disclosure due to their openness over shared networks. Research on privacy protection of MASs is thus of great significance (Nozari et al., 2017; Gao L et al., 2019; Kishida, 2019; Ruan et al., 2019; Wang YQ, 2019; Yin et al., 2020; Fang et al., 2021; Wang

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 62473204 and 61873130), the "Chunhui Program" Collaborative Scientific Research Project, China (No. 202202004), the Natural Science Foundation of Nanjing University of Posts and Telecommunications, China (Nos. NY221082, NY222144, and NY223075), and the Huali Program for Excellent Talents in Nanjing University of Posts and Telecommunications, China

Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/FITEE.2300532) contains supplementary materials, which are available to authorized users

ORCID: Yang YANG, https://orcid.org/0000-0002-8706-2831; Dong YUE, https://orcid.org/0000-0001-7810-9338

YQ et al., 2021; Xu et al., 2021; Liang et al., 2022; Chen et al., 2023).

Many techniques have been developed for preserving privacy in consensus control, such as homomorphic encryption (Fang et al., 2021; Wang YQ et al., 2021; Chen et al., 2023), differential privacy (Yang ZW et al., 2022), and state decomposition (Wang YQ, 2019). Specifically, Chen et al. (2023) addressed a privacy-preserving economic dispatch issue for microgrids by a homomorphically encrypted consensus strategy. Yan et al. (2021) and Wang AJ et al. (2022) offered distributed privacy protection algorithms and consensus-based economic dispatch strategies, ensuring information security and node privacy. In Fang et al. (2021), a homomorphic encryption based protocol was generalized for a second-order system. In Wang YQ et al. (2021), an agent decomposition strategy was introduced via Paillier encryption, and the security of agents' initial states was analyzed. It is worth pointing out that these results do not consider the impact of privacy-preserving errors in an MAS, and these errors are of paramount importance in consensus behaviors.

Reinforcement learning (RL) plays a vital role in consensus control. A common architecture of RL is a critic–actor form. With the actor–critic architecture, state-feedback and output-feedback optimal control methods were proposed in Tong et al. (2018) and Li YM et al. (2020a, 2022) under the framework of backstepping technique. Li YM et al. (2020a) investigated an adaptive fuzzy inverse optimal control problem and developed an inverse optimal scheme. To address the issue of adaptive control of large-scale systems in strict-feedback form, a feedforward decentralized controller with adaptive laws was designed in Tong et al. (2018) using a state observer with the backstepping design technique. For strict-feedback systems containing unknown internal dynamics and immeasurable constrained states, an adaptive neural network output-feedback optimized control algorithm was designed in Li YM et al. (2022). For the optimal distributed consensus control of discrete-time MASs with unknown dynamics, Yang XD et al. (2022) proposed a data-based distributed control algorithm to enhance online learning capabilities using policy gradient RL. For discrete-time MASs with dead zones, an adaptive fault-tolerant tracking scheme (Li HY et al., 2021) was developed by combining the backstepping technique with an RL algorithm. For the consensus problem in nonlinear MASs, an off-policy model-free algorithm based on RL (Wang H and Li, 2022) was proposed for achieving fully cooperative consensus, thereby further expanding the scope of the research. Subsequently, an optimization leader-following consensus control method based on RL (Wen and Li, 2022) was developed for a second-order MAS with unknown dynamics. To address the consensus problem of higher-order nonlinear MASs with uncertainties and communication delays, a sliding mode control design method (Li JN et al., 2023) was developed based on the principle of sliding mode control and RL techniques.

In spite of the fruitful results in consensus control for MASs, there still exist several technical challenges:

1. It is noteworthy that, although the aforementioned consensus results (Li HY et al., 2021; Wang H and Li, 2022; Wen and Li, 2022; Yang XD et al., 2022; Li JN et al., 2023) are from RL techniques, they fail to consider privacy leakage risks during information transmission among agents, where state information is sensitive in some application scenarios.

2. In Liu DX (2013), the Liu cryptosystem uses a large number of random sequences to create a ciphertext. Actually, although these random numbers increase data privacy, they lead to a decryption error between a plaintext and a decrypted plaintext. There is still a lack of theory for improving decryption performance. Further technique is thus necessary to mitigate decryption errors arising from random numbers.

3. Since there inevitably exist decryption errors in a cryptosystem, given that the existence of errors poses a technical challenge for consensus analysis, we may infer that such errors have an impact on consensus performance. Regrettably, the existing results (Wang YQ, 2019; Fang et al., 2021; Wang YQ et al., 2021; Liang et al., 2022) on encryption techniques lack analysis for decryption errors, let alone elucidations of the means for dealing with these errors.

To fill in these gaps, we propose an RL-based privacy-preserving consensus tracking control strategy for a nonstrict-feedback discrete MAS. The specific technical contributions of this paper are summarized below.

1. An improved Liu cryptosystem is proposed via signal amplification. Signal amplification is for

reduction of the error between encryption and decryption. Different from the Paillier encryption algorithm (Gao C et al., 2021; Yang ZW et al., 2022), an amplification operation is performed on plaintext before the encryption process in this research, increasing its proportion in the ciphertext. According to the decryption algorithm, the decrypted plaintext contains two parts, namely the original plaintext and a random component. The amplification of the plaintext results in a reduction in the proportion of the random component. As a result, the impact of the error between encryption and decryption on the plaintext is alleviated, allowing for satisfactory recovery of the plaintext information.

2. An RL-based privacy-preserving consensus tracking control strategy is proposed for nonstrict-feedback discrete-time MASs. With the improved privacy-preserving mechanism, an RL technology is introduced. Different from RL methods (Li HY et al., 2021; Wang H and Li, 2022; Wen and Li, 2022; Yang XD et al., 2022; Li JN et al., 2023), the strategy safeguards sensitive information to ensure the privacy of agents' interaction. For error signals generated by the improved privacy-preserving mechanism, our RL strategy treats them as unknown terms and compensates them online. This strategy not only provides effective solutions for protecting sensitive information but also improves consensus performance for an MAS.

2 Problem formulation and preliminaries

2.1 Graph theory

A graph is introduced to describe information flow among agents. A directed graph $\mathcal{G}=\{\mathcal{V},\mathcal{E},\boldsymbol{\mathcal{A}}\}$ is considered, where $\mathcal{V}=\{1,2,\cdots,N\}$ represents a set of non-empty nodes, $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ is the edge set that fulfills the condition, and $\boldsymbol{\mathcal{A}}=\left[a_{i,j}\right]\in\mathbb{R}^{N\times N}$ represents the connectivity among agents. Edge $(j,i)\in\mathcal{E}$ stands for an edge from node j to node i. In $\boldsymbol{\mathcal{A}}$, $a_{i,j}>0$ if node i is able to send information to node j. Otherwise, $a_{i,j}=0$. $a_{i,i}=0$ denotes the prohibition of self-transmission and self-reception of information. $\boldsymbol{\mathcal{D}}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_N\}$, among which $d_i=\sum_{j=1}^{N}a_{i,j}$, is a diagonal matrix. The Laplacian matrix of $\mathcal{G}$ is $\boldsymbol{\mathcal{L}}=\boldsymbol{\mathcal{D}}-\boldsymbol{\mathcal{A}}$. The neighbor set of the i^{th} agent is $\mathcal{N}_i=\{j|(j,i)\in\mathcal{E}\}$. $\bar{\mathcal{G}}=\{\bar{\mathcal{V}},\bar{\mathcal{E}}\}$ is an augmented graph, where $\bar{\mathcal{V}}=\{0,1,\cdots,N\}$ and $\bar{\mathcal{E}}=\bar{\mathcal{V}}\times\bar{\mathcal{V}}$ with 0 being the leader. If node i can receive the information from the leader, $a_{i,0}>0$; otherwise, $a_{i,0}=0$. We define $\boldsymbol{\mathcal{B}}=\mathrm{diag}\{a_{1,0},a_{2,0},\cdots,a_{N,0}\}$.

2.2 Problem formulation

We consider a nonstrict-feedback MAS consisting of one leader and N followers with an n^{th}-order form, and the dynamics of the i^{th} follower is

$$\begin{cases}x_{i,1}(k+1)=f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))+g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))x_{i,2}(k),\\ x_{i,2}(k+1)=f_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k))+g_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k))x_{i,3}(k),\\ \qquad\vdots\\ x_{i,n_i}(k+1)=f_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))+g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))u_i(k)\\ \qquad\qquad+\varpi_i(k),\\ y_i(k)=x_{i,1}(k),\end{cases}\tag{1}$$

where $i=1,2,\cdots,N$, $x_{i,l}(k)\in\mathbb{R}$ is a state variable with $l=1,2,\cdots,n_i$, $\bar{\boldsymbol{x}}_{i,n_i}(k)=[x_{i,1}(k),x_{i,2}(k),\cdots,x_{i,n_i}(k)]^{\mathrm{T}}\in\mathbb{R}^{n_i}$ represents the state vector, $y_i(k)\in\mathbb{R}$ is the follower's output, $u_i(k)\in\mathbb{R}$ is an input signal, $f_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k))$ is an unknown and smooth function denoting uncertain dynamics, $g_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k))$ is a non-zero function representing control gains, and $\varpi_i(k)$ denotes external disturbance, which is unknown but bounded, satisfying $|\varpi_i(k)|\leq\bar{\varpi}_i$ with $\bar{\varpi}_i\geq 0$.

Remark 1 The distinctive feature of nonstrict feedback is that the unknown nonlinear functions $f_{i,l}(\cdot)$ and $g_{i,l}(\cdot)$ of each subsystem include all state variables, implying that all state variables exist in the l^{th}-order differential subsystem of the i^{th} follower (Sun KK et al., 2021; Sun JL et al., 2022). Their applications are also widespread, such as marine surface vessels (Qi et al., 2023), wheeled mobile robots (Ding et al., 2021), single-link robots (Zhou et al., 2018), and electromechanical systems (Li YM et al., 2020b). However, nonstrict-feedback followers still pose challenges in control law design due to the occurrence of the algebraic loop problem (Tong et al., 2016) when applying the existing adaptive backstepping technique. Therefore, it is necessary to develop new design approaches.

Assumption 1 (Ding et al., 2020, 2021) The function $g_{i,l}(\cdot)$ is unknown but bounded, satisfying $0<\underline{g}_{i,l}\leq|g_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k))|\leq\bar{g}_{i,l}$, where $\underline{g}_{i,l}$ and $\bar{g}_{i,l}$ are unknown positive constants.

Assumption 2 The topology graph $\bar{\mathcal{G}}$ contains a spanning tree with the leader being the root node.

We define the consensus tracking error $\boldsymbol{e} = \boldsymbol{y} - (y_{\mathrm{d}} \otimes \boldsymbol{I}_N)$ between the leader and the followers, where y_{d} is the leader's signal, $\boldsymbol{y} = [y_1, y_2, \cdots, y_N]^{\mathrm{T}}$, "$\otimes$" is the Kronecker product, and $\boldsymbol{I}_N$ is an N-dimensional identity vector. If the norm of this error converges to a small neighborhood around the origin, the MAS is said to achieve consensus tracking.

In a complex network environment, there exists a vulnerability for the exchange of sensitive information to take place in a way that reaches malicious eavesdroppers or adversaries, resulting in adverse effects on the consensus performance. This paper focuses mainly on addressing the privacy-preserving consensus tracking issue. Followers perform tracking control according to the leader's signal and output signals from their neighbors. The information transmission process is shown in Fig. 1. The output signal y_i in the MAS and y_{d} from the leader are exposed to public networks and can be easily intercepted, potentially compromising the tracking consensus. It is thus necessary to investigate methods for ensuring the security of the transmission information, such as y_{d} and y_i, thereby guaranteeing a privacy-preserving consensus.

The control objective of this paper is to develop an RL-based privacy-preserving control strategy that enables an MAS to achieve consensus tracking while ensuring that all signals are ultimately bounded.

For a follower in Eq. (1), a distributed synchronization error for follower i is defined as

$$\begin{aligned} e_{i,1}(k) = &\sum_{j\in\mathcal{N}_i} a_{i,j}(y_i(k) - y_j(k)) \\ &+ a_{i,0}(y_i(k) - y_{\mathrm{d}}(k)), \end{aligned} \tag{2}$$

and a vector containing all synchronization errors is denoted as $\boldsymbol{e}_1 = [e_{1,1}, e_{2,1}, \cdots, e_{N,1}]^{\mathrm{T}}$. According to the knowledge related to the graph theory, we obtain

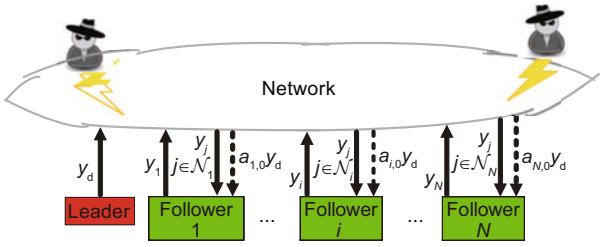

Fig. 1 Information transmission process in a multi-agent system

the relationship between $\boldsymbol{e}_1$ and $\boldsymbol{e}$ as follows:

$$\boldsymbol{e}_1 = (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})\boldsymbol{e} = \boldsymbol{\mathcal{H}}\boldsymbol{e}. \tag{3}$$

3 Reinforcement learning (RL) based strategy with a privacy-preserving method

In this section, an RL-based adaptive consensus tracking control strategy with a privacy-preserving method is proposed for a nonstrict-feedback nonlinear discrete-time MAS. This control strategy incorporates an improved Liu cryptosystem, critic neural networks (NNs), and actor NNs. The improved Liu cryptosystem is used primarily for encryption and decryption of followers' information. Critic NNs use synchronization errors as reward and penalty signals to design the utility function, while actor NNs combine Liu decrypted information with the utility function to design control laws. More details are presented in the following subsections.

3.1 Improved Liu cryptosystem

One of the common approaches adopted for addressing the privacy-preserving issue presented in Section 2.2 is the safeguarding of information via encrypted transmission. Currently, a series of encryption techniques appear, such as Rivest–Shamir–Adleman (RSA) (Rivest et al., 1978) and Elgamal (Elgamal, 1985) algorithms. In this paper, we introduce a cryptosystem called the Liu cryptosystem, and it holds distinct characteristics compared with the aforementioned algorithms. Liu cryptosystem contains a private-key-based algorithm that employs only private keys during the encryption and decryption processes without the requirement for a public key. In contrast to asymmetric encryption algorithms, which involve the utilization of both public and private keys, the Liu cryptosystem adopts a simpler encryption method while achieving an equivalent level of confidentiality.

Remark 2 Compared with RSA (Reddy et al., 2023) and homomorphic encryption (Zuo et al., 2021; Zhang P et al., 2023), the encryption method in this paper requires only a single key for encryption and decryption, with simplified key management. This cryptosystem reduces requirements for data format, allowing for direct encryption and decryption of diverse data, whereas the methods in other studies

(Gao C et al., 2021; Yang ZW et al., 2022; Chen et al., 2023) require data to be converted into integers, unavoidably introducing quantization errors. Compared with the traditional Liu cryptosystem (Liu DX, 2013), the improved method yields the result that decrypted plaintext is closer to the original one under identical conditions.

The Liu cryptosystem (Liu DX, 2013) is employed to encrypt the information exchanged among the MASs. Its principle is to generate a private key consisting of four random numbers and share it with the sender and receiver. Then the sender encrypts the plaintext using the private key for transmission, and the receiver decrypts the ciphertext using the same private key for obtaining the plaintext. The Liu cryptosystem consists of two parts, namely the encryption and decryption algorithms. The block diagram of the Liu cryptosystem is shown in Fig. 2. In this block diagram, a private key $K(M)=\{(q_1,s_1,t_1),(q_2,s_2,t_2),\cdots,(q_M,s_M,t_M)\}$ is generated, where $q_i\in\mathbb{R}$, $s_i\in\mathbb{R}$, and $t_i\in\mathbb{R}$ are randomly generated and shared with the sender and receiver, $i=1,2,\cdots,M$, and $M\in\mathbb{Z}$. This cryptosystem requires $M\geq 3$, $q_j\neq 0$, and $q_M+s_M+t_M\neq 0$, and only one j satisfies $t_j\neq 0$, $1\leq j\leq M-1$. Upon receiving the private key $K(M)$, the sender uses the Liu encryption algorithm $\mathrm{Enc}(K(M),m)$ to encrypt the plaintext m, resulting in a series of ciphertext $\bar{c}_{m,M}=\{c_1,c_2,\cdots,c_M\}$, which is transmitted over the network. At the receiver's terminal, upon receiving the ciphertext $\bar{c}_{m,M}$, the Liu decryption algorithm $\mathrm{Dec}(K(M),\bar{c}_{m,M})$ is applied using the private key $K(M)$ to obtain the decrypted plaintext $\hat{m}$. With this transmission process, sensitive messages over networks are protected.

Based on a commonsense perspective, the Liu cryptosystem has been designed in a way that uses a large number of random numbers to create a ciphertext. According to the results in Liu DX (2013), the creation of ciphertext is from $c_i=q_it_im+s_ir_M+q_i(r_i-r_{i-1})$ in the Liu encryption algorithm. In the Liu decryption algorithm, $\hat{m}=(\sum_{i=1}^{M-1}(c_i-Ss_i)/q_i)/T=(\sum_{i=1}^{M-1}(\frac{A_i}{q_i}+\frac{B_i-Ss_i}{q_i}))/T$, where $T=\sum_{i=1}^{M-1}t_i$, and $S=\frac{c_M}{q_M+t_M+s_M}$. We further obtain $\hat{m}=m+\varepsilon$, where $\varepsilon=(\sum_{i=1}^{M-1}\frac{B_i-Ss_i}{q_i})/T$. This indicates that ε determines the decryption performance. To elaborate, if the value of $\frac{\varepsilon}{m+\varepsilon}$ becomes smaller, the relationship between $\hat{m}$ and m reveals that $\hat{m}$ is closer to m. There exist two methods for reduction of the value of $\frac{\varepsilon}{m+\varepsilon}$: the first is to decrease ε, and the second is to amplify the value of m. Actually, in the first method, ε contains numerous random numbers, and it is difficult to adjust its size precisely. We prefer the second method and are going to amplify the value of m for improvement of the decryption performance.

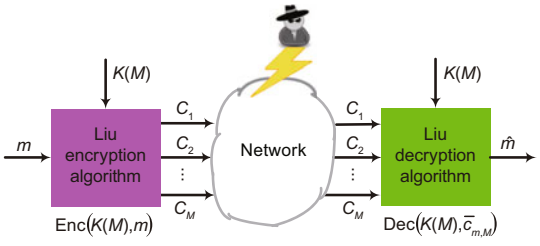

Fig. 2 Block diagram of the Liu cryptosystem

From the phenomenon and motivation, an improved Liu cryptosystem is proposed. Its main idea is to increase the magnitude of plaintext and reduce the magnitude of random terms in ciphertext calculation, and the action alleviates the impact of random terms on the error between m and $\hat{m}$, resulting in a better reproducibility of m with privacy protection. Its process is summarized in Algorithm 1, and its block diagram is shown in Fig. 3.

Remark 3 As the magnitude of the data is sufficiently large, it is not necessary to multiply it by k_p before encryption actually takes place. However, due to the inability to accurately determine the magnitude of the data, it is a common practice to multiply the data by an amplification parameter k_p before performing the encryption operation. From

Algorithm 1 Improved Liu cryptosystem

1: Initialization of parameters M, k_p, $T=0$, $T_1=0$;
2: Plaintext $\rightarrow m$
3: Amplify the signal $m_\mathrm{A}=k_\mathrm{p}m$
4: **Generation of the private key** $K(M)$: randomly generate r_n, q_n, t_n, and s_n, $n=1,2,\cdots,M$;
5: **Encryption**$(\mathrm{Enc}(K(M),m_\mathrm{A}))$: $c_1=q_1t_1m_\mathrm{A}+s_1r_M+q_1(r_1-r_{M-1})$;
6: $i=2$ to $M-1$
$c_i=q_it_im_\mathrm{A}+s_ir_M+q_i(r_i-r_{i-1})$, $c_M=(q_Ms_Mt_M)r_M$, $c_{M+1}=c_M$;
7: **Decryption**$(\mathrm{Dec}(K(M),\bar{c}_{m_\mathrm{A},M}))$
8: $i=2$ to $M-1$
$T=T+t_i$,
$S=c_M/(q_M+s_M+t_M)$;
9: $i=2$ to $M-1$
$T_1=T_1+(c_i-Ss_i)/q_i$,
$\hat{m}_\mathrm{A}=T_1/T$
10: **Output** $\hat{m}=\hat{m}_\mathrm{A}/k_\mathrm{p}$

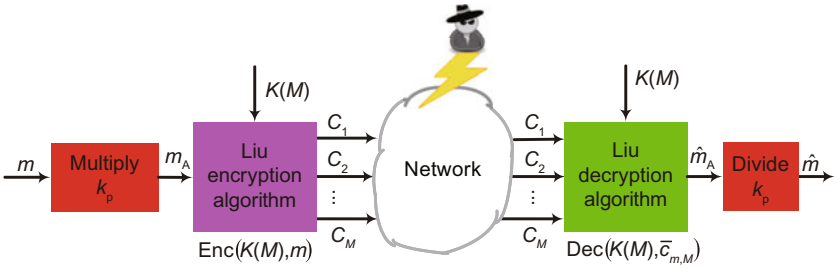

Fig. 3 Block diagram of the improved Liu cryptosystem

$c_i = q_i t_i m_{\mathrm{A}} + s_i r_M + q_i(r_i - r_{i-1})$ and $\hat{m} = m + \varepsilon$ with $\varepsilon = \frac{1}{T}(\sum_{i=1}^{M-1} \frac{B_i - S s_i}{q_i})$, the analysis indicates that the privacy information is the transformation between random numbers and plaintext.

The gain k_{p} serves primarily to amplify plaintext into $m_{\mathrm{A}} = k_{\mathrm{p}} m$, and this operation reduces the value of $\frac{\varepsilon}{\varepsilon + m}$ for describing the impact of encryption and decryption errors. Output information is encrypted into ciphertext for transmission with k_{p} being independent of this process. It is broadcasted to all agents but is not considered encrypted content. Even if k_{p} leaks, eavesdroppers lacking decryption algorithms cannot access any of the actual information. Thus, the leakage of k_{p} does not exercise a negative impact on privacy preservation.

3.2 Strategic utility function and critic neural networks

For a follower in information reception, it is necessary to decrypt an encrypted signal for obtaining the true value. The distributed synchronization error $e_{i,1}(k)$ in Eq. (2) is thus represented as $\breve{e}_{i,1}(k) = \sum_{j \in \mathcal{N}_i} a_{i,j}(y_i(k) - \hat{y}_j(k)) + a_{i,0}(y_i(k) - \hat{y}_{\mathrm{d}}(k))$, where $\hat{y}_i(k)$ is the decrypted information of follower i via the improved Liu decryption algorithm, and $\hat{y}_{\mathrm{d}}(k)$ is the decrypted information of the leader.

According to Li HY et al. (2021) and Bai et al. (2020b), the utility function is chosen as

$$\zeta_i(k) = \begin{cases} 0, & \text{if } |\breve{e}_{i,1}(k)| \leq \sigma, \\ 1, & \text{if } |\breve{e}_{i,1}(k)| > \sigma, \end{cases} \tag{4}$$

where σ is a threshold related to the tolerance of consensus tracking performance, and $\zeta_i(k)$ is a real-time index depending on the consensus tracking performance. With the utility function (4), we consider the long-term performance of the MAS and define the long-term strategic utility function as

$$\phi_i(k) = \beta_i^{\rho} \zeta_i(k+1) + \beta_i^{\rho-1} \zeta_i(k+2) + \cdots + \beta_i^{k+1} \zeta_i(\rho), \tag{5}$$

where β_i is a weight satisfying $0 < \beta_i < 1$, ρ is the total number of steps, and $\phi_i(k)$ represents the cumulative performance index with subsequent steps.

With the critic NNs approximating the long-term strategic utility function (5), we obtain $\phi_i(k) = (\boldsymbol{\theta}_{i,\phi}^*)^{\mathrm{T}} \boldsymbol{S}_{i,\phi}(k) + \varepsilon_{i,\phi}(k)$, where $\boldsymbol{\theta}_{i,\phi}^* = \left[\theta_{i,\phi,1}^*, \theta_{i,\phi,2}^*, \cdots, \theta_{i,\phi,\gamma_{i,\phi}}^*\right]^{\mathrm{T}} \in \mathbb{R}^{\gamma_{i,\phi}}$ is the ideal weight vector, $\gamma_{i,\phi}$ is the number of NN's nodes, $\boldsymbol{S}_{i,\phi}(k)$ is the NN basis function vector, and $\varepsilon_{i,\phi}(k)$ is the approximate error. From Bai et al. (2020b), we perceive that $\|\boldsymbol{\theta}_{i,\phi}^*\| \leq \bar{\theta}_{i,\phi}$ and $|\varepsilon_{i,\phi}(k)| \leq \bar{\varepsilon}_{i,\phi}$ with $\bar{\theta}_{i,\phi}$ and $\bar{\varepsilon}_{i,\phi}$ being unknown positive constants. In the following part related to the design of control strategy, these properties are covered at each step.

As $\phi_i(k)$ is an unknown ideal value, only its estimated value $\hat{\phi}_i(k) = \hat{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k) \boldsymbol{S}_{i,\phi}(k)$ can be obtained, where $\hat{\boldsymbol{\theta}}_{i,\phi}$ is the estimate of $\boldsymbol{\theta}_{i,\phi}$. $E_{i,\phi}(k) = \beta_i \hat{\phi}_i(k) - \left[\hat{\phi}_i(k-1) - \zeta_i(k)\right]$ represents the Bellman error. Based on $J_{i,\phi}(k) = \frac{1}{2} E_{i,\phi}^2(k)$ and the method from Bai et al. (2020b), the derivation of $\hat{\boldsymbol{\theta}}_{i,\phi}(k)$ can take as

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{i,\phi}(k+1) = \hat{\boldsymbol{\theta}}_{i,\phi}(k) - \mu_{i,\phi} \beta_i \boldsymbol{S}_{i,\phi}(k) \big[\beta_i \hat{\phi}_i(k) \\ - \hat{\phi}_i(k-1) + \zeta_i(k)\big], \end{aligned} \tag{6}$$

where $\mu_{i,\phi}$ is a positive parameter.

3.3 RL-based privacy-preserving consensus tracking control strategy

Similar to transformation technology for a nonlinear discrete system in Ge SS et al. (2003) and Bai et al. (2020a), for facilitating the strategy design, the

i^{th} follower (1) is transformed as

$$\begin{cases} x_{i,1}(k+n_i) = f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k)) \\ \qquad\qquad +g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))x_{i,2}(k+n_i-1), \\ x_{i,2}(k+n_i-1) = f_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k)) \\ \qquad\qquad +g_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k))x_{i,3}(k+n_i-2), \\ x_{i,n_i}(k+1) = f_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k)) \\ \qquad\qquad +g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))u_i(k)+\varpi_i(k), \\ y_i(k) = x_{i,1}(k),\ i=1,2,\cdots,N. \end{cases} \tag{7}$$

In what follows, an RL-based adaptive consensus tracking control strategy will be given for Eq. (7).

Step 1: From the definition of $e_{i,1}$ in Eq. (2) and the dynamics of the follower, $e_{i,1}(k+n_i) = \sum_{j\in\mathcal{N}_i} a_{i,j}(y_i(k+n_i)-y_j(k+n_i)) + a_{i,0}(y_i(k+n_i)-y_{\text{d}}(k+n_i))$. We define $e_{i,2}(k+n_i-1) = x_{i,2}(k+n_i-1)-\alpha_{i,2}(k)$, where $\alpha_{i,2}(k)$ is the virtual control law at this step, and will be determined later. From Eq. (7), we obtain

$$\begin{aligned} &e_{i,1}(k+n_i) \\ =&(d_i+a_{i,0})g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1)) \\ &\cdot\Big[\Big(\frac{f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} \\ &-\frac{v_iy_j(k+n_i)+(1-v_i)y_{\text{d}}(k+n_i)}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} \\ &+v_iy_j(k+n_i)+(1-v_i)y_{\text{d}}(k+n_i)\Big) \\ &-v_iy_j(k+n_i)-(1-v_i)y_{\text{d}}(k+n_i) \\ &+\alpha_{i,2}(k)+e_{i,2}(k+n_i-1)\Big], \end{aligned} \tag{8}$$

where $v_i = \frac{d_i}{d_i+a_{i,0}}$. We denote $\psi_{i,1}(k) = -\frac{f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} + \frac{v_iy_j(k+n_i)+(1-v_i)y_{\text{d}}(k+n_i)}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} - v_iy_j(k+n_i)-(1-v_i)y_{\text{d}}(k+n_i)$.

According to the results in Section 3.1, for better information security, the transmitted signals are encrypted by the improved Liu encryption algorithm. Upon receiving the ciphertext, the signal is decrypted using the improved Liu decryption algorithm to obtain the desired message. There exists the following relationship between the true signal and the decrypted one:

$$\hat{z}_i(k) = y_i(k)-\hat{y}_i(k), \tag{9}$$

$$\hat{z}_{\text{d}}(k) = y_{\text{d}}(k)-\hat{y}_{\text{d}}(k), \tag{10}$$

where $y_{\text{d}}(k)$ is the output information of the leader, $\hat{z}_i(k)$ is the error between the output information y_i of the i^{th} follower and a decrypted information $\hat{y}_i$, and $\hat{z}_{\text{d}}(k)$ is the error between the leader's output information y_{d} and the decrypted one $\hat{y}_{\text{d}}$. In fact, $\hat{z}_i(k)$ and $\hat{z}_{\text{d}}(k)$ are unknown and unpredictable for follower j. These errors are treated as unknown variables and are composed into unknown functions. According to Eqs. (9) and (10), $\psi_{i,1}(k)$ is rewritten as

$$\begin{aligned} \psi_{i,1}(k) =& -\frac{f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} \\ &+\frac{v_i(\hat{y}_j(k+n_i)+\hat{z}_j(k+n_i))}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} \\ &+\frac{(1-v_i)(\hat{y}_{\text{d}}(k+n_i)+\hat{z}_{\text{d}}(k+n_i))}{g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1))} \\ &-v_i(\hat{y}_j(k+n_i)+\hat{z}_j(k+n_i)) \\ &-(1-v_i)(\hat{y}_{\text{d}}(k+n_i)+\hat{z}_{\text{d}}(k+n_i)). \end{aligned} \tag{11}$$

Due to $\psi_{i,1}(k)$ being an ideal value, it cannot be obtained. Using an actor NN to obtain the estimated value $\psi_{i,1}(k) = (\boldsymbol{\theta}_{i,1}^*)^{\text{T}}\boldsymbol{S}_{i,1}(\boldsymbol{Z}_{i,1}(k))+\varepsilon_{i,1}(k)$, where $\boldsymbol{\theta}_{i,1}^* \in \mathbb{R}^{\gamma_{i,1}}$ is the ideal weight vector, $\gamma_{i,1}$ is the number of NN's nodes, $\boldsymbol{S}_{i,1}(\cdot)$ is the NN basis function vector, $\varepsilon_{i,1}(k)$ is the approximate error, and $\boldsymbol{Z}_{i,1}(k) = \big[x_{i,1}(k+n_i-1), x_{i,2}(k+n_i-1),\cdots,x_{i,n_i}(k+n_i-1), \sum_{j\in\mathcal{N}_i}\hat{y}_j(k+n_i), \hat{y}_{\text{d}}(k+n_i)\big]^{\text{T}}$. Then, error (8) becomes

$$\begin{aligned} &e_{i,1}(k+n_i) \\ &= (d_i+a_{i,0})g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-1)) \\ &\cdot\Big[-(\boldsymbol{\theta}_{i,1}^*)^{\text{T}}\boldsymbol{S}_{i,1}(\boldsymbol{Z}_{i,1}(k))-\varepsilon_{i,1}(k) \\ &+\alpha_{i,2}(k)+(\boldsymbol{\theta}_{i,1}^*)^{\text{T}}\boldsymbol{S}_{i,1}(\hbar_{i,1}(k)) \\ &-(\boldsymbol{\theta}_{i,1}^*)^{\text{T}}\boldsymbol{S}_{i,1}(\hbar_{i,1}(k))+e_{i,2}(k+n_i-1) \\ &-v_i(\hat{y}_j(k+n_i)+\hat{z}_j(k+n_i))-(1-v_i)(\hat{y}_d(k+n_i) \\ &+\hat{z}_d(k+n_i))\Big], \end{aligned} \tag{12}$$

where $\hbar_{i,1}(k) = \big[x_{i,1}(k), \sum_{j\in\mathcal{N}_i}\hat{y}_j(k), \hat{y}_{\text{d}}(k+n_i)\big]^{\text{T}}$.

We design the virtual control law

$$\begin{aligned} \alpha_{i,2}(k) =& \hat{\boldsymbol{\theta}}_{i,1}^{\text{T}}(k)\boldsymbol{S}_{i,1}(\hbar_{i,1}(k))+v_i\hat{y}_j(k) \\ &+(1-v_i)\hat{y}_{\text{d}}(k+n_i) \end{aligned} \tag{13}$$

with an adaptive update law

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{i,1}(k+1) =& \hat{\boldsymbol{\theta}}_{i,1}(k_{i,1})-\mu_{i,1}\boldsymbol{S}_{i,1}(\hbar_{i,1}(k_{i,1})) \\ &\cdot\big[\hat{\boldsymbol{\theta}}_{i,1}^{\text{T}}(k_{i,1})\boldsymbol{S}_{i,1}(\hbar_{i,1}(k_{i,1}))+\hat{\phi}_i(k)\big], \end{aligned} \tag{14}$$

where $\mu_{i,1}$ is a positive parameter, $k_{i,1}=k-n_i+1$, $\hat{\boldsymbol{\theta}}_{i,1}$ is the estimate of $\boldsymbol{\theta}^*_{i,1}$, and $\hat{\phi}_i(k)$ has been reported in Section 3.2.

Step l ($2\le l\le n_i-1$): We define the l^{th} error as $e_{i,l}(k+n_i+1-l)=x_{i,l}(k+n_i+1-l)-\alpha_{i,l}(k)$, where $\alpha_{i,l}(k)$ is the virtual control law at step $l-1$.

With the dynamics of the follower in Eq. (1), we similarly denote $\psi_{i,l}(k)=-\frac{f_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-l))}{g_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-l))}+\frac{\alpha_{i,l}(k)}{g_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-l))}-\alpha_{i,l}(k)$. We employ the NN approximation technique $\psi_{i,l}(k)=(\boldsymbol{\theta}^*_{i,l})^{\mathrm{T}}\boldsymbol{S}_{i,l}(\boldsymbol{Z}_{i,l}(k))+\varepsilon_{i,l}(k)$, where $\boldsymbol{Z}_{i,l}(k)=\big[x_{i,1}(k+n_i-l),x_{i,2}(k+n_i-l),\cdots,x_{i,n_i}(k+n_i-l),\sum_{j\in\mathcal{N}_i}\hat{y}_j(k+n_i),\hat{y}_{\mathrm{d}}(k+n_i)\big]^{\mathrm{T}}$. From the signal analysis here, we substitute information from neighbors and/or the leader with the signals generated from the improved Liu cryptosystem instead of the real-time values in Bai et al. (2020b). This is one of key differences from the results in Bai et al. (2020b), and it arises owing to the requirement for privacy preservation. Then, the l^{th} error yields

$$\begin{aligned}&e_{i,l}(k+n_i+1-l)\\=&g_{i,l}(\bar{\boldsymbol{x}}_{i,n_i}(k+n_i-l))\big[-(\boldsymbol{\theta}^*_{i,l})^{\mathrm{T}}\boldsymbol{S}_{i,l}(\boldsymbol{Z}_{i,l}(k))\\&+(\boldsymbol{\theta}^*_{i,l})^{\mathrm{T}}\boldsymbol{S}_{i,l}(\hbar_{i,l}(k))+e_{i,l+1}(k+n_i-l)\\&-(\boldsymbol{\theta}^*_{i,l})^{\mathrm{T}}\boldsymbol{S}_{i,l}(\hbar_{i,l}(k))-\varepsilon_{i,l}(k)+\alpha_{i,l+1}(k)-\alpha_{i,l}(k)\big],\end{aligned}\tag{15}$$

where $\hbar_{i,l}(k)=\big[x_{i,1}(k),x_{i,2}(k),\cdots,x_{i,l}(k),\sum_{j\in\mathcal{N}_i}\hat{y}_j(k),\hat{y}_{\mathrm{d}}(k+n_i)\big]^{\mathrm{T}}$, $e_{i,l+1}(k+n_i-l)=x_{i,l+1}(k+n_i-l)-\alpha_{i,l+1}(k)$ is the $(l+1)^{\text{th}}$ error, and $\alpha_{i,l+1}(k)$ is the virtual control law at step l.

For stabilizing $e_{i,l}$, we design a virtual control law at this step

$$\alpha_{i,l+1}(k)=\hat{\boldsymbol{\theta}}^{\mathrm{T}}_{i,l}(k)\boldsymbol{S}_{i,l}(\hbar_{i,l}(k))+\alpha_{i,l}(k),\tag{16}$$

with an adaptive update law

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{i,l}(k+1)=&\hat{\boldsymbol{\theta}}_{i,l}(k_{i,l})-\mu_{i,l}\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_{i,l}))\\&\cdot\big[\hat{\boldsymbol{\theta}}^{\mathrm{T}}_{i,l}(k_{i,l})\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_{i,l}))+\hat{\phi}_i(k)\big],\end{aligned}\tag{17}$$

where $\mu_{i,l}$ is a positive parameter, $k_{i,l}=k-n_i+l$, and $\hat{\boldsymbol{\theta}}_{i,l}$ is the estimate of $\boldsymbol{\theta}^*_{i,l}$.

Step n_i: We define the n_i^{th} error variable $e_{i,n_i}(k+1)=x_{i,n_i}(k+1)-\alpha_{i,n_i}(k)$. This error becomes

$$\begin{aligned}e_{i,n_i}(k+1)=&g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))\Big[\Big(\frac{f_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}{g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}\\&-\frac{\alpha_{i,n_i-1}(k)}{g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}+\alpha_{i,n_i-1}(k)\Big)\\&+u_i(k)-\alpha_{i,n_i}(k)\Big]+\varpi_i(k),\end{aligned}\tag{18}$$

where $\alpha_{i,n_i}(k)$ is the virtual control law at step n_i-1.

We denote an unknown function $\psi_{i,n_i}(k)=-\frac{f_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}{g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}+\frac{\alpha_{i,n_i}(k)}{g_{i,n_i}(\bar{\boldsymbol{x}}_{i,n_i}(k))}-\alpha_{i,n_i}(k)$, and an actor NN is employed to approximate this unknown function $\psi_{i,n_i}(k)=(\boldsymbol{\theta}^*_{i,n_i})^{\mathrm{T}}\boldsymbol{S}_{i,n_i}(\boldsymbol{Z}_{i,n_i}(k))+\varepsilon_{i,n_i}(k)$, where $\boldsymbol{Z}_{i,n_i}(k)=\big[x_{i,1}(k),x_{i,2}(k),\cdots,x_{i,n_i}(k),\sum_{j\in\mathcal{N}_i}\hat{y}_j(k),\hat{y}_{\mathrm{d}}(k+n_i)\big]^{\mathrm{T}}$.

Then the final control law is designed:

$$\begin{aligned}u_i(k)=&-\breve{c}_{i,1}\breve{e}_{i,1}(k)-\sum_{l=2}^{n_i}\breve{c}_{i,l}e_{i,l}(k)\\&+\hat{\boldsymbol{\theta}}^{\mathrm{T}}_{i,n_i}(k)\boldsymbol{S}_{i,n_i}(\hbar_{i,n_i}(k))+\alpha_{i,n_i}(k),\end{aligned}\tag{19}$$

with an adaptive update law

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{i,n_i}(k+1)=&\hat{\boldsymbol{\theta}}_{i,n_i}(k_{n_i})-\mu_{i,n_i}\boldsymbol{S}_{i,n_i}(\hbar_{i,n_i}(k_{n_i}))\\&\cdot\big[\hat{\boldsymbol{\theta}}^{\mathrm{T}}_{i,n_i}(k_{n_i})\boldsymbol{S}_{i,n_i}(\hbar_{i,n_i}(k_{n_i}))+\hat{\phi}_i(k)\big],\end{aligned}\tag{20}$$

where $\breve{c}_{i,l}$ is a control gain, μ_{i,n_i} is a positive parameter, $k_{n_i}=k$, $\hat{\boldsymbol{\theta}}_{i,n_i}$ is the estimate of $\boldsymbol{\theta}^*_{i,n_i}$, and $\hbar_{i,n_i}(k)=\boldsymbol{Z}_{i,n_i}(k)$.

The block diagram of the overall MAS is shown in Fig. 4, where the output signals of the leader and followers are connected through a directed communication topology, and the strategy is constructed with the use of privacy-preserving information received from neighboring followers. In the MAS, the encryption and decryption processes are for transmission information, critic NNs use synchronization errors as a reward and punishment for a utility function, and actor NNs use approximation of a utility function for u_i recursively. The diagram of the proposed strategy for the i^{th} follower is shown in Fig. 5.

Remark 4 The dynamics of the i^{th} follower (1) comprises n_i subsystems. There exists an issue in selecting virtual control laws $\alpha_{i,2}(k)=-[f_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))-y_{\mathrm{d}}(k+1)]/g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))$ and $\alpha_{i,3}(k)=-[f_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k))-\alpha_{i,2}(k+1)]/g_{i,2}(\bar{\boldsymbol{x}}_{i,n_i}(k))$ for stabilization. While $\alpha_{i,2}(k)$ stabilizes the first subsystem, $\alpha_{i,3}(k)$ stabilizes the second subsystem. However, $\alpha_{i,2}(k+1)$ represents a future virtual control law

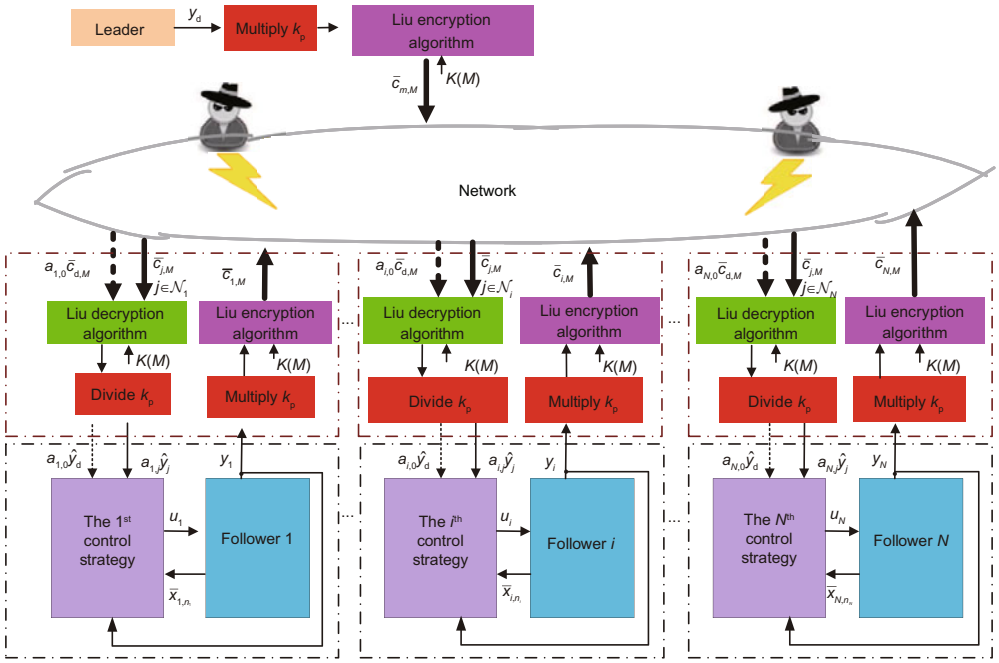

Fig. 4 Block diagram of the multi-agent system (MAS)

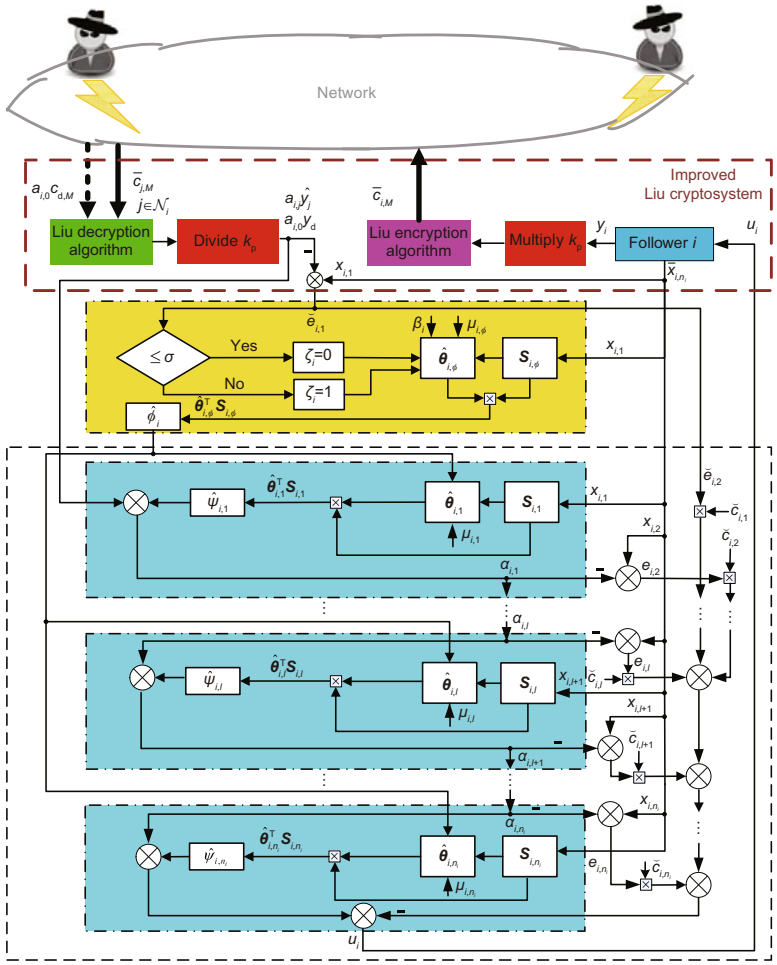

Fig. 5 Block diagram of the i^{th} follower

that is not yet available, and $\alpha_{i,3}(k)$ is infeasible. To address this issue, the dynamics of the i^{th} follower is transformed into a specific form of Eq. (7). The transformation allows us to treat the original follower as a one-step predictor, transforming it into an equivalent maximum n_i-step predictor. Thereafter, $\alpha_{i,3}(k)$, as well as the virtual control signals arising in pursuance, becomes feasible. More details can be found in Ge SS et al. (2003).

Remark 5 In our strategy, an improved Liu cryptosystem is introduced, which encrypts the information transmitted among agents and preserves privacy. The differences between our strategy and the one based on the existing RL method in Bai et al. (2020b) can be stated from the points of view of the following three perspectives: the construction of the utility function, the analysis of error dynamics, and the design of control laws. The control strategy based on the existing RL method uses real information, while our strategy employs decrypted information for privacy preservation.

4 Stability analysis

Here, the stability is analyzed. The main results are summarized in the following theorem:

Theorem 1 For a nonstrict-feedback nonlinear MAS with the leader and followers (1) under Assumption 1, the RL-based adaptive consensus tracking control strategy with a privacy-preserving method consists of the virtual control laws (13) and (16), the parameter update laws (6), (14), (17), and (20), and the control law (19). If the design parameters satisfy $0<\beta_i<1$, $0<\mu_{i,\phi}<1/(\beta_i^2\gamma_{i,\phi})$, $0<\mu_{i,l}<1/\gamma_{i,l}$, and $\breve{c}_{i,l}>0$, the RL-based control strategy with the privacy-preserving method ensures that all signals in the closed-loop system are ultimately bounded and that the consensus tracking error of the MAS achieves consensus tracking.

Proof Substituting Eqs. (13), (16), and (19) into Eqs. (12), (15), and (18) at instant $k+1$, respectively, we have the following error dynamics:

$$\begin{aligned}
&e_{i,1}(k+1)\\
=&(d_i+a_{i,0})g_{i,1}(\bar{x}_{i,n_i}(k))\Big[-\boldsymbol{\theta}_{i,1}^{\mathrm{T}}(k_{i,1})\boldsymbol{S}_{i,1}(\boldsymbol{Z}_{i,1}(k_{i,1}))\\
&+\boldsymbol{\theta}_{i,1}^{\mathrm{T}}(k_{i,1})\boldsymbol{S}_{i,1}(\boldsymbol{\hbar}_{i,1}(k_{i,1}))-\varepsilon_{i,1}(k_{i,1})\\
&+e_{i,2}(k)-\tilde{\boldsymbol{\theta}}_{i,1}^{\mathrm{T}}(k_{i,1})\boldsymbol{S}_{i,1}(\boldsymbol{\hbar}_{i,1}(k_{i,1}))\Big],
\end{aligned}\tag{21}$$

$$\begin{aligned}
&e_{i,l}(k+1)\\
=&g_{i,1}(\bar{\boldsymbol{x}}_{i,n_i}(k))\Big[-\boldsymbol{\theta}_{i,l}^{\mathrm{T}}(k_{i,l})\boldsymbol{S}_{i,l}(\boldsymbol{Z}_{i,l}(k_{i,l}))\\
&-\varepsilon_{i,l}(k_{i,l})+\boldsymbol{\theta}_{i,l}^{\mathrm{T}}(k_{i,l})\boldsymbol{S}_{i,l}(\boldsymbol{\hbar}_{l}(k_{i,l}))\\
&-\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_{i,l})\boldsymbol{S}_{i,l}(\boldsymbol{\hbar}_{l}(k_{i,l}))+e_{i,l}(k)\Big],
\end{aligned}\tag{22}$$

$$\begin{aligned}
e_{i,n_i}(k+1)=&g_{i,n}(\bar{\boldsymbol{x}}_{i,n_i}(k))\Big[\tilde{\boldsymbol{\theta}}_{i,n_i}^{\mathrm{T}}(k)\boldsymbol{S}_{i,n_i}(\boldsymbol{\hbar}_{i,n_i}(k))\\
&-\varepsilon_{i,n_i}(k)-\sum_{l=1}^{n_i}\breve{c}_{i,l}e_{i,l}\Big]+\varpi_i(k),
\end{aligned}\tag{23}$$

where $\tilde{\boldsymbol{\theta}}_{i,1}(k)=\hat{\boldsymbol{\theta}}_{i,1}(k)-\boldsymbol{\theta}_{i,1}^{*}$, $\tilde{\boldsymbol{\theta}}_{i,l}(k)=\hat{\boldsymbol{\theta}}_{i,l}(k)-\boldsymbol{\theta}_{i,l}^{*}$, and $\tilde{\boldsymbol{\theta}}_{i,n_i}(k)=\hat{\boldsymbol{\theta}}_{i,n_i}(k)-\boldsymbol{\theta}_{i,n_i}^{*}$.

Consider a Lyapunov function as follows:

$$V(k)=V_1(k)+V_2(k),\tag{24}$$

where $V_1(k)=\frac{\omega_{e,1}}{4}e_{i,1}^2(k)+\sum_{l=2}^{n_i-1}\frac{\omega_{e,l}}{4}e_{i,l}^2(k)+\frac{\omega_{e,n_i}}{n_i+3}e_{i,n_i}^2(k)$, $V_2(k)=\sum_{l=1}^{n_i}\frac{\omega_{\theta,l}}{\mu_{i,l}}\sum_{\epsilon=0}^{n_i-l}\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_l+\epsilon)\tilde{\boldsymbol{\theta}}_{i,l}(k_l+\epsilon)+\frac{\omega_{\phi}}{\mu_{i,\phi}}\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\tilde{\boldsymbol{\theta}}_{i,\phi}(k)+6\omega_{\phi}\big[\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k-1)\boldsymbol{S}_{i,\phi}(k-1)\big]^2$, $\omega_{\phi}>0$, $\omega_{\theta,l}>0$, and $\omega_{e,l}>0$.

With Young's inequality and the inequality $0<\boldsymbol{S}_{i,l}^{\mathrm{T}}(\cdot)\boldsymbol{S}_{i,l}(\cdot)\le\gamma_{i,l}$ from the property in the radial basis function (RBF) NN, we obtain

$$-\boldsymbol{\theta}_{i,l}^{\mathrm{T}}\boldsymbol{S}_{i,l}(\boldsymbol{Z}_{i,l}(k_l))+\boldsymbol{\theta}_{i,l}^{\mathrm{T}}\boldsymbol{S}_{i,l}(\boldsymbol{\hbar}_{i,l}(k_l))\le\bar{\boldsymbol{\theta}}_{i,l}^2+\gamma_{i,l},\tag{25}$$

where $l=1,2,\cdots,n_i-1$.

By Eq. (21) and inequality (25) and the Cauchy–Schwarz inequality $(\chi_1+\chi_2+\cdots+\chi_n)^2\le n(\chi_1^2+\chi_2^2+\cdots+\chi_n^2)$, where $\chi_m\in\mathbb{R}$, $m=1,2,\cdots,n$, we deduce the forward difference of $V_1(k)$ as

$$\begin{aligned}
&\triangle V_1(k)\\
\le&(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2\Big(\tilde{\boldsymbol{\theta}}_{i,1}^{\mathrm{T}}(k_{i,1})\boldsymbol{S}_{i,1}(\boldsymbol{\hbar}_{i,1}(k_{i,1}))\Big)^2\\
&+(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2e_{i,2}^2(k)+(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2\bar{\varepsilon}_{i,1}^2\\
&+(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2(\bar{\boldsymbol{\theta}}_{i,1}^2+\gamma_{i,l})^2-\frac{\omega_{e,1}}{4}e_{i,1}^2(k)\\
&+\sum_{l=2}^{n_i-1}\omega_{e,l}\bar{g}_{i,l}^2\Big(\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_{i,l})\boldsymbol{S}_{i,l}(\boldsymbol{\hbar}_{i,l}(k_{i,l}))\Big)^2\\
&+\sum_{l=2}^{n_i-1}\omega_{e,l}\bar{g}_{i,l}^2e_{i,l+1}^2(k)+\sum_{l=2}^{n_i-1}\omega_{e,l}\bar{g}_{i,l}^2\bar{\varepsilon}_{i,l}^2\\
&+\sum_{l=2}^{n_i-1}\omega_{e,l}\bar{g}_{i,l}^2(\bar{\boldsymbol{\theta}}_{i,l}^2+\gamma_{i,l})^2-\sum_{l=2}^{n_i-1}\frac{\omega_{e,l}}{4}e_{i,l}^2(k)
\end{aligned}$$

$$+\omega_{e,n_i}\bar{g}_{i,n_i}^2\Big(\tilde{\boldsymbol{\theta}}_{i,n_i}^{\mathrm{T}}(k_{n_i})\boldsymbol{S}_{i,n_i}(\hbar_{i,n_i}(k_{n_i}))\Big)^2+\omega_{e,n_i}\bar{g}_{i,n_i}^2\sum_{l=1}^{n_i}\breve{c}_{i,l}^2e_{i,l}^2-\frac{\omega_{e,n_i}}{n+3}e_{i,n_i}^2(k)+\omega_{e,n_i}\bar{g}_{i,n_i}^2\bar{\varepsilon}_{i,n_i}^2+\omega_{e,n_i}\bar{\varpi}_i^2. \tag{26}$$

From Eqs. (14), (17), and (20), combining the Cauchy–Schwarz inequality and the inequality $0<\boldsymbol{S}_{i,\phi}^{\mathrm{T}}(\cdot)\boldsymbol{S}_{i,\phi}(\cdot)\le\gamma_{i,\phi}$, the forward difference of $V_2(k)$ is obtained as

$$\begin{aligned}&\triangle V_2(k)\\ \le&-\sum_{l=1}^{n_i}\omega_{\theta,l}(1-\mu_{i,l}\gamma_{i,l})\big(\hat{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_l)\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_l))\\&+\hat{\phi}_i(k)\big)^2-\sum_{l=1}^{n_i}\omega_{\theta,l}\Big(\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_l)\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_l))\Big)^2\\&+\sum_{l=1}^{n_i}2\omega_{\theta,l}\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\boldsymbol{S}_{i,\phi}(k)\Big)^2\\&+\sum_{l=1}^{n_i}2\omega_{\theta,l}\Big(\frac{1}{2}\bar{\boldsymbol{\theta}}_{i,l}^2+\gamma_{i,l}+\frac{1}{2}\bar{\boldsymbol{\theta}}_{i,\phi}^2\Big)^2\\&-\omega_\phi(1-\mu_{i,\phi}\beta_i^2\gamma_{i,\phi})\Big(\beta_i\hat{\phi}_i(k)-\hat{\phi}_i(k-1)\\&+\zeta_i(k)\Big)^2+6\omega_\phi\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k-1)\boldsymbol{S}_{i,\phi}(k-1)\Big)^2+3\omega_\phi\\&+\frac{3}{4}\omega_\phi(2+\beta_i^2)\Big(\bar{\boldsymbol{\theta}}_{i,\phi}^2+\gamma_{i,\phi}\Big)^2-\omega_\phi\beta_i^2\\&\cdot\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\boldsymbol{S}_{i,\phi}(k)\Big)^2+6\omega_\phi\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\boldsymbol{S}_{i,\phi}(k)\Big)^2\\&-6\omega_\phi\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k-1)\boldsymbol{S}_{i,\phi}(k-1)\Big)^2.\end{aligned} \tag{27}$$

From inequalities (26) and (27), by selecting the parameters $\mu_{i,l}$ and $\mu_{i,\phi}$ satisfying $0<\mu_{i,l}\le1/\gamma_{i,l}$ and $0<\mu_{i,\phi}\le1/(\beta_i^2\gamma_{i,\phi})$ respectively, we have

$$\begin{aligned}&\triangle V(k)\\ \le&-\sum_{l=1}^{n_i-2}\Big(\frac{\omega_{e,l+1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,l+1}^2-\omega_{e,l}\bar{g}_{i,l}^2\Big)e_{i,l+1}^2(k)\\&-\Big(\frac{\omega_{e,1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,1}^2\Big)e_{i,1}^2(k)-\Big(\frac{\omega_{e,n_i}}{n+3}\\&-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,n_i}^2-\omega_{e,n_i-1}\bar{g}_{i,n_i-1}^2\Big)e_{i,n_i}^2\\&-\Big(\omega_\phi\beta_i^2-\sum_{l=1}^{2}2\omega_{\theta,l}-6\omega_\phi\Big)\Big(\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\boldsymbol{S}_{i,\phi}(k)\Big)^2\\&-\sum_{l=1}^{n_i}(\omega_{\theta,l}-\omega_{e,l}\bar{g}_{i,l}^2)\Big(\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_{i,l})\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_{i,l}))\Big)^2+D,\end{aligned} \tag{28}$$

where $D=\sum_{l=1}^{n_i}2\omega_{\theta,l}\Big(\frac{1}{2}\bar{\boldsymbol{\theta}}_{i,l}^2+\gamma_{i,l}+\frac{1}{2}\bar{\boldsymbol{\theta}}_{i,\phi}^2\Big)^2+3\omega_\phi+\frac{3}{4}\omega_\phi(2+\beta_i^2)\Big(\bar{\boldsymbol{\theta}}_{i,\phi}^2+\gamma_{i,\phi}\Big)^2+\omega_{e,n_i}\bar{\varpi}_i^2+\sum_{l=2}^{n_i-1}\omega_{e,l}\bar{g}_{i,l}^2(\bar{\boldsymbol{\theta}}_{i,l}^2+\gamma_{i,l})^2+(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2\bar{\varepsilon}_{i,1}^2+(d_i+a_{i,0})^2\omega_{e,1}\bar{g}_{i,1}^2(\bar{\boldsymbol{\theta}}_{i,1}^2+\gamma_{i,l})^2+\sum_{l=2}^{n_i}\omega_{e,l}\bar{g}_{i,l}^2\bar{\varepsilon}_{i,l}^2$.

We choose parameters such as $\frac{\omega_{e,l+1}}{4}>\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,l+1}^2+\omega_{e,l}\bar{g}_{i,l}^2$, $l=1,2,\cdots,n_i-2$, $\frac{\omega_{e,n_i}}{n_i+3}>\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,n_i}^2+\omega_{e,n_i-1}\bar{g}_{i,n_i-1}^2$, $\omega_\phi\beta_i^2>\sum_{l=1}^{2}2\omega_{\theta,l}+6\omega_\phi$, $\omega_{\theta,l}>\omega_{e,l}\bar{g}_{i,l}^2$, $\frac{\omega_{e,1}}{4}>\omega_{e,n_i}\bar{g}_{i,n_i}^2c_{i,1}^2$. It is concluded that $\triangle V(k)<0$ as long as $|e_{i,l+1}(k)|>\frac{\sqrt{D}}{\sqrt{\frac{\omega_{e,l+1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,l+1}^2-\omega_{e,l}\bar{g}_{i,l}^2}}$ ($l=1,2,\cdots,n_i-2$), $|e_{i,n_i}(k)|>\frac{\sqrt{D}}{\sqrt{\frac{\omega_{e,n_i}}{n_i+3}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,n_i}^2-\omega_{e,n_i-1}\bar{g}_{i,n_i-1}^2}}$, $|e_{i,1}(k)|>\frac{\sqrt{D}}{\sqrt{\frac{\omega_{e,1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,1}^2}}$, $|\tilde{\boldsymbol{\theta}}_{i,\phi}^{\mathrm{T}}(k)\boldsymbol{S}_{i,\phi}(k)|>\frac{\sqrt{D}}{\sqrt{\omega_\phi\beta_i^2-\sum_{l=1}^{2}2\omega_{\theta,l}-6\omega_\phi}}$, and $|\tilde{\boldsymbol{\theta}}_{i,l}^{\mathrm{T}}(k_l)\boldsymbol{S}_{i,l}(\hbar_{i,l}(k_l))|>\frac{\sqrt{D}}{\sqrt{\omega_{\theta,l}-\omega_{e,l}\bar{g}_{i,l}^2}}$, $l=1,2,\cdots,n_i$.

Now we are going to present the bound for consensus tracking error in the MAS. From the stability analysis, $e_{i,1}(k)$ converges to $\frac{\sqrt{D}}{\sqrt{\frac{\omega_{e,1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,1}^2}}$. According to Eq. (3) and Assumption 2, we obtain $\boldsymbol{e}=\boldsymbol{\mathcal{H}}^{-1}\boldsymbol{e}_1$ (Zhang HG et al., 2017). Then, the norm of the output consensus tracking error $\|\boldsymbol{e}\|$ converges to $\sqrt{\sum_{i=1}^{N}\frac{\sqrt{D}}{\sqrt{\frac{\omega_{e,1}}{4}-\omega_{e,n_i}\bar{g}_{i,n_i}^2\breve{c}_{i,1}^2}}}\Big/\sigma_{\min}(\boldsymbol{\mathcal{H}})$, where $\sigma_{\min}(\boldsymbol{\mathcal{H}})$ is the minimum singular value of $\boldsymbol{\mathcal{H}}$. The errors in the MAS as well as other signals in the closed-loop system are ultimately bounded. The proof is completed.

The guideline for the selection of parameters in the RL-based control strategy with a privacy-preserving method is given as follows:

1. Select parameters $0<\sigma<1$, $0<\beta_i<1$, and $0<\mu_{i,\phi}<1/(\beta_i^2\gamma_{i,\phi})$. Determine the utility function (4), the long-term policy utility function (5), and the update law $\hat{\boldsymbol{\theta}}_{i,\phi}(k)$.

2. Choose the parameter $0<\mu_{i,l}<1/\gamma_{i,l}$ and determine the virtual control laws $\alpha_{i,l}(k)$ with update laws for $\hat{\boldsymbol{\theta}}_{i,l}(k)$ ($l=1,2,\cdots,n_i$).

3. By selecting control parameters $\breve{c}_{i,l}>0$ ($l=1,2,\cdots,n_i$) and $0<\mu_{i,n_i}<1/\gamma_{i,n_i}$, the control law $u_i(k)$ with an update law for $\hat{\boldsymbol{\theta}}_{i,n_i}(k)$ is determined.

5 Simulation examples

To demonstrate the feasibility of the theoretical claims, two simulation examples are given. In this

section, the two MASs both consist of one leader and four followers, and there exists the risk of privacy disclosure over communication topology. Fig. 6 shows the communication topology.

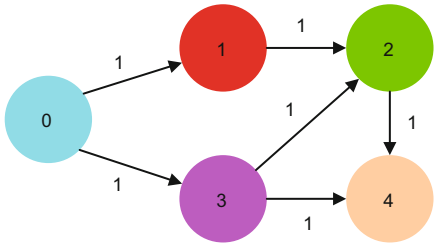

Fig. 6 Communication graph in Examples 1 and 2

Example 1 Consider the four followers in Eq. (1):

$$\begin{cases} x_{i,1}(k+1) = f_{i,1}(\bar{\boldsymbol{x}}_{i,2}(k)) + g_{i,1}(\bar{\boldsymbol{x}}_{i,2}(k))x_{i,2}(k), \\ x_{i,2}(k+1)=f_{i,2}(\bar{\boldsymbol{x}}_{i,2}(k))+g_{i,2}(\bar{\boldsymbol{x}}_{i,2}(k))u_i(k)+\varpi_i(k), \\ y_i(k) = x_{i,1}(k), \end{cases} \tag{29}$$

where $f_{i,1}(\bar{\boldsymbol{x}}_{i,2}(k)) = -x_{i,1}(k) \triangle T$, $f_{i,2}(\bar{\boldsymbol{x}}_{i,2}(k)) = \big[-x_{i,2}(k) + 3.1(0.4 - x_{i,1}(k))\exp\big(\frac{1.5x_{i,2}(k)}{3.4+x_{i,2}(k)}\big)\big] \triangle T$, $g_{i,1}(\bar{\boldsymbol{x}}_{i,2}(k)) = \frac{1}{x_{i,2}(k)}\big[0.019(1.5+0.01x_{i,1}^2(k))\exp\big(\frac{4x_{i,2}(k)}{3.4+x_{i,2}(k)}\big)\big] \triangle T$, $g_{i,2}(\bar{\boldsymbol{x}}_{i,2}(k)) = 4.1(1.001 + \sin(x_{i,1}(k)x_{i,2}(k))) \triangle T$, $\varpi_i(k) = 0.05\cos(0.05k)\cos(x_{i,1}(k))$, and $\triangle T$ is the sampling time.

The reference signal is given by $y_{\mathrm{d}} = 0.006\sin(\pi/8+0.6\pi t/38)+0.0365$. The centers of the RBF NN for $\phi_i(k)$, $\psi_{i,1}(k)$, and $\psi_{i,2}(k)$ are uniformly spaced in $[-2\times2]\times[-2\times2]\times[-2\times2]$ with width $\eta = \sqrt{2}$. The index threshold is chosen as $\sigma = 0.0002$. The initial state values are $x_{1,1}(0) = 0.09$, $x_{2,1}(0) = 0.095$, $x_{3,1}(0) = 0.12$, $x_{4,1}(0) = 0.1$, $x_{i,2}(0) = 0.5$, $\hat{\boldsymbol{\theta}}_{i,\phi}(0) = [0.003, 0.003, \cdots, 0.003]^{\mathrm{T}}_{50\times1}$, $\hat{\boldsymbol{\theta}}_{i,1}(0) = [0.0003, 0.0003, \cdots, 0.0003]^{\mathrm{T}}_{50\times1}$, and $\hat{\boldsymbol{\theta}}_{i,2}(0) = [0.0003, 0.0003, \cdots, 0.0003]^{\mathrm{T}}_{50\times1}$. Other control parameters are $\beta_i = 0.00001$, $\mu_{i,\phi} = 0.02$, $\mu_{i,1} = 0.5$, $\mu_{i,2} = 0.2$, $\check{c}_{i,1} = 120$, and $\check{c}_{i,2} = 0.2$.

Due to the sensitivity of the output information from both leaders and followers, it is necessary to encrypt this information before communication interaction among agents. The simulation results are displayed in Figs. 7–11. The curves of the leader and followers are plotted in Fig. 7. The control inputs

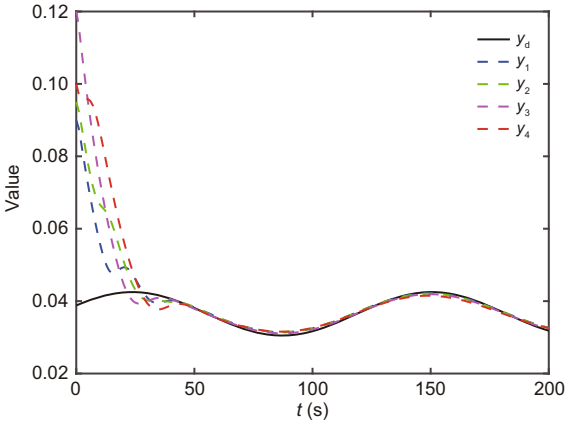

Fig. 7 y_i $(i = 1, 2, 3, 4)$ and y_{d} in Example 1

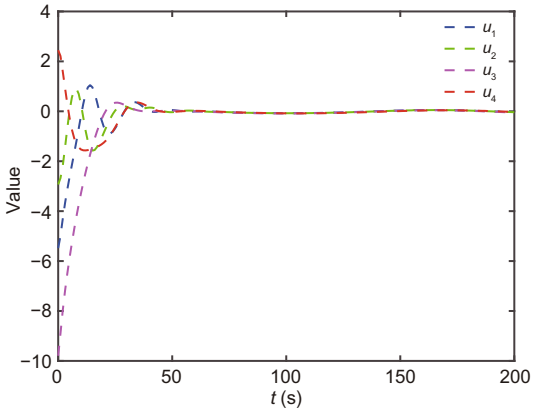

Fig. 8 Control inputs in Example 1

are plotted in Fig. 8. In Fig. 9, the output information of follower 4 is presented, which is similar with ciphertext after being decrypted by the improved Liu decryption algorithm, and the encrypted information during the transmission process is also provided. From Fig. 10, we observe that the decryption error associated with our improved Liu cryptosystem is smaller than that arising pursuant to using the algorithm from Liu DX (2013). For an intuitive comparison, we take the leader as an example. As seen from Fig. 11, the amplified plaintext is able to better restore the truth value after decryption.

Example 2 Marine surface vessels are taken as another example.

The details of Example 2 are provided in the supplementary materials.

6 Conclusions

An RL-based privacy-preserving consensus tracking control strategy with improved Liu

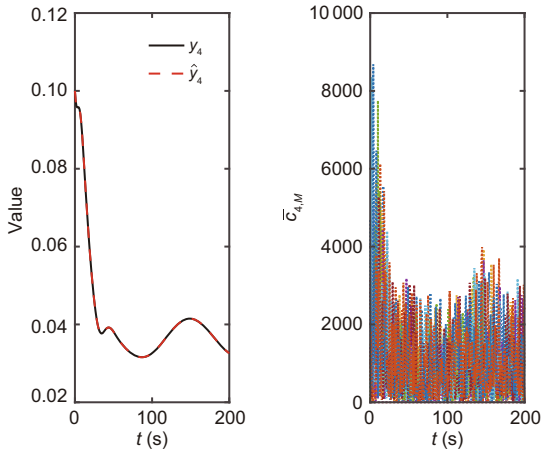

Fig. 9 Relevant information and ciphertext of follower 4 in Example 1

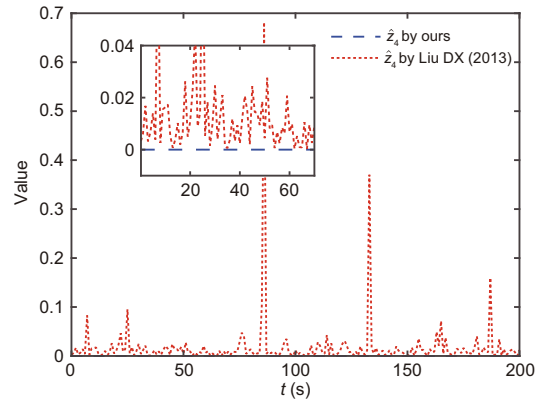

Fig. 10 Comparison of follower 4's decryption errors with different methods in Example 1

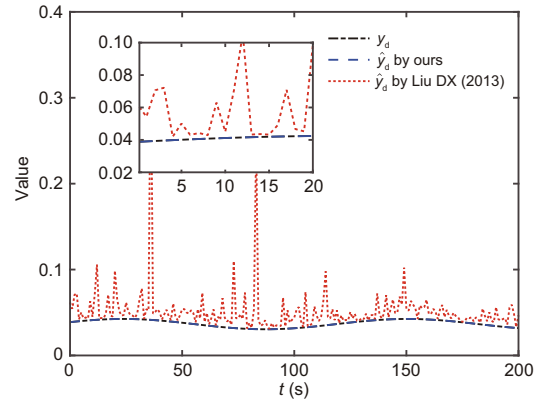

Fig. 11 Comparison of decryption signals of the leader with different methods in Example 1

cryptosystem has been proposed for a nonstrict-feedback discrete-time MAS. We observe that the improved Liu cryptosystem encrypts the information transmitted among agents and thus effectively prevents the possibility of intrusion of external information. With the help of a Lyapunov function, the proposed control strategy has guaranteed that all signals are ultimately bounded and that the output consensus errors of the MAS converge to a neighborhood around the origin. This control strategy is only for nonstrict-feedback MASs, and it is not applicable to non-affine ones (Wang SB, 2022), MASs with state (Liu L et al., 2022), output constraints (Liu L et al., 2021; Lin et al., 2022), or input saturation (Zhang JX et al., 2021). Therefore, our further work will focus on general MASs with both state and output constraints, concomitant with fulfilling the priority of privacy preservation.

Contributors

Yang YANG, Fanming HUANG, and Dong YUE designed the research. Yang YANG and Fanming HUANG processed the data and drafted the paper. Dong YUE helped organize the paper. Yang YANG and Fanming HUANG revised and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Bai WW, Zhang B, Zhou Q, et al., 2020a. Multigradient recursive reinforcement learning NN control for affine nonlinear systems with unmodeled dynamics. *Int J Robust Nonl Contr*, 30(4):1643-1663. https://doi.org/10.1002/rnc.4843

Bai WW, Li TS, Tong SC, 2020b. NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems. *IEEE Trans Cybern*, 50(11):4573-4584. https://doi.org/10.1109/TCYB.2020.2963849

Chen W, Liu L, Liu GP, 2023. Privacy-preserving distributed economic dispatch of microgrids: a dynamic quantization-based consensus scheme with homomorphic encryption. *IEEE Trans Smart Grid*, 14(1):701-713. https://doi.org/10.1109/TSG.2022.3189665

Ding L, Li S, Gao HB, et al., 2020. Adaptive partial reinforcement learning neural network-based tracking control for wheeled mobile robotic systems. *IEEE Trans Syst Man Cybern Syst*, 50(7):2512-2523. https://doi.org/10.1109/TSMC.2018.2819191

Ding L, Li S, Gao HB, et al., 2021. Adaptive neural network-based finite-time online optimal tracking control of the nonlinear system with dead zone. *IEEE Trans Cybern*,

51(1):382-392.
https://doi.org/10.1109/TCYB.2019.2939424

Elgamal T, 1985. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Trans Inform Theory*, 31(4):469-472.
https://doi.org/10.1109/TIT.1985.1057074

Fang WT, Zamani M, Chen ZY, 2021. Secure and privacy preserving consensus for second-order systems based on Paillier encryption. *Syst Contr Lett*, 148:104869.
https://doi.org/10.1016/j.sysconle.2020.104869

Gao C, Wang ZD, He X, et al., 2021. Encryption–decryption-based consensus control for multi-agent systems: handling actuator faults. *Automatica*, 134:109908.
https://doi.org/10.1016/j.automatica.2021.109908

Gao L, Deng SJ, Ren W, 2019. Differentially private consensus with an event-triggered mechanism. *IEEE Trans Contr Netw Syst*, 6(1):60-71.
https://doi.org/10.1109/TCNS.2018.2795703

Ge SS, Li GY, Lee TH, 2003. Adaptive NN control for a class of strict-feedback discrete-time nonlinear systems. *Automatica*, 39(5):807-819.
https://doi.org/10.1016/S0005-1098(03)00032-3

Ge XH, Xiao SY, Han QL, et al., 2022. Dynamic event-triggered scheduling and platooning control co-design for automated vehicles over vehicular ad-hoc networks. *IEEE/CAA J Autom Sin*, 9(1):31-46.
https://doi.org/10.1109/JAS.2021.1004060

Ge XH, Han QL, Wu Q, et al., 2023. Resilient and safe platooning control of connected automated vehicles against intermittent denial-of-service attacks. *IEEE/CAA J Autom Sin*, 10(5):1234-1251.
https://doi.org/10.1109/JAS.2022.105845

Ge XH, Han QL, Zhang XM, et al., 2024. Communication resource-efficient vehicle platooning control with various spacing policies. *IEEE/CAA J Autom Sin*, 11(2):362-376. https://doi.org/10.1109/JAS.2023.123507

He WL, Xu WY, Ge XH, et al., 2022. Secure control of multiagent systems against malicious attacks: a brief survey. *IEEE Trans Ind Inform*, 18(6):3595-3608.
https://doi.org/10.1109/TII.2021.3126644

Ju YM, Ding DR, He X, et al., 2022. Consensus control of multi-agent systems using fault-estimation-in-the-loop: dynamic event-triggered case. *IEEE/CAA J Autom Sin*, 9(8):1440-1451.
https://doi.org/10.1109/JAS.2021.1004386

Kishida M, 2019. Encrypted control system with quantiser. *IET Contr Theory Appl*, 13(1):146-151.
https://doi.org/10.1049/iet-cta.2018.5764

Li HY, Wu Y, Chen M, 2021. Adaptive fault-tolerant tracking control for discrete-time multiagent systems via reinforcement learning algorithm. *IEEE Trans Cybern*, 51(3):1163-1174.
https://doi.org/10.1109/TCYB.2020.2982168

Li JN, Yuan L, Chai TY, et al., 2023. Consensus of nonlinear multiagent systems with uncertainties using reinforcement learning based sliding mode control. *IEEE Trans Circ Syst I Regular Papers*, 70(1):424-434.
https://doi.org/10.1109/TCSI.2022.3206102

Li P, Hu JP, Qiu L, et al., 2022. A distributed economic dispatch strategy for power-water networks. *IEEE Trans Contr Netw Syst*, 9(1):356-366.
https://doi.org/10.1109/TCNS.2021.3104103

Li YM, Min X, Tong SC, 2020a. Adaptive fuzzy inverse optimal control for uncertain strict-feedback nonlinear systems. *IEEE Trans Fuzzy Syst*, 28(10):2363-2374.
https://doi.org/10.1109/TFUZZ.2019.2935693

Li YM, Shao XF, Tong SC, 2020b. Adaptive fuzzy prescribed performance control of nontriangular structure nonlinear systems. *IEEE Trans Fuzzy Syst*, 28(10):2416-2426.
https://doi.org/10.1109/TFUZZ.2019.2937046

Li YM, Liu YJ, Tong SC, 2022. Observer-based neuro-adaptive optimized control of strict-feedback nonlinear systems with state constraints. *IEEE Trans Neur Netw Learn Syst*, 33(7):3131-3145.
https://doi.org/10.1109/TNNLS.2021.3051030

Liang CD, Ge MF, Xu JZ, et al., 2022. Secure and privacy-preserving formation control for networked marine surface vehicles with sampled-data interactions. *IEEE Trans Veh Technol*, 71(2):1307-1318.
https://doi.org/10.1109/TVT.2021.3133902

Lin XZ, Chen CC, Li SH, 2022. Finite-time output feedback stabilization for a class of output-constrained planar switched systems. *IEEE Trans Circ Syst II Express Briefs*, 69(1):164-168.
https://doi.org/10.1109/TCSII.2021.3077603

Liu DX, 2013. Homomorphic Encryption for Database Querying. US Patent 20150295716.

Liu L, Ding SH, Yu XH, 2021. Second-order sliding mode control design subject to an asymmetric output constraint. *IEEE Trans Circ Syst II Express Briefs*, 68(4):1278-1282. https://doi.org/10.1109/TCSII.2020.3021715

Liu L, Cui YJ, Liu YJ, et al., 2022. Adaptive event-triggered output feedback control for nonlinear switched systems based on full state constraints. *IEEE Trans Circ Syst II Express Briefs*, 69(9):3779-3783.
https://doi.org/10.1109/TCSII.2022.3173679

Ning BD, Han QL, Zuo ZY, et al., 2023. Fixed-time and prescribed-time consensus control of multiagent systems and its applications: a survey of recent trends and methodologies. *IEEE Trans Ind Inform*, 19(2):1121-1135. https://doi.org/10.1109/TII.2022.3201589

Nozari E, Tallapragada P, Cortés J, 2017. Differentially private average consensus: obstructions, trade-offs, and optimal algorithm design. *Automatica*, 81:221-231.
https://doi.org/10.1016/j.automatica.2017.03.016

Peng ZH, Wang J, Wang D, et al., 2021. An overview of recent advances in coordinated control of multiple autonomous surface vehicles. *IEEE Trans Ind Inform*, 17(2):732-745.
https://doi.org/10.1109/TII.2020.3004343

Qi XJ, Liu WH, Lu Y, 2023. Event-triggered-based fuzzy adaptive tracking control for nonstrict-feedback asymmetric state constrained systems. *Fuzzy Sets Syst*, 470:108642.
https://doi.org/10.1016/j.fss.2023.108642

Reddy SS, Sinha S, Zhang W, 2023. Design and analysis of RSA and Paillier homomorphic cryptosystems using PSO-based evolutionary computation. *IEEE Trans*

Comput, 72(7):1886-1900.
https://doi.org/10.1109/TC.2023.3234213

Rivest RL, Shamir A, Adleman L, 1978. A method for obtaining digital signatures and public-key cryptosystems. *Commun ACM*, 21(2):120-126.
https://doi.org/10.1145/359340.359342

Ruan MH, Gao H, Wang YQ, 2019. Secure and privacy-preserving consensus. *IEEE Trans Autom Contr*, 64(10):4035-4049.
https://doi.org/10.1109/TAC.2019.2890887

Sakthivel R, Sakthivel R, Kaviarasan B, et al., 2019. Finite-time leaderless consensus of uncertain multi-agent systems against time-varying actuator faults. *Neurocomputing*, 325:159-171.
https://doi.org/10.1016/j.neucom.2018.10.020

Shahvali M, Naghibi-Sistani MB, Askari J, 2018. Adaptive output-feedback bipartite consensus for nonstrict-feedback nonlinear multi-agent systems: a finite-time approach. *Neurocomputing*, 318:7-17.
https://doi.org/10.1016/j.neucom.2018.07.039

Sun JL, Yi JQ, Pu ZQ, 2022. Fixed-time adaptive fuzzy control for uncertain nonstrict-feedback systems with time-varying constraints and input saturations. *IEEE Trans Fuzzy Syst*, 30(4):1114-1128.
https://doi.org/10.1109/TFUZZ.2021.3052610

Sun KK, Qiu JB, Karimi HR, et al., 2021. A novel finite-time control for nonstrict feedback saturated nonlinear systems with tracking error constraint. *IEEE Trans Syst Man Cybern Syst*, 51(6):3968-3979.
https://doi.org/10.1109/TSMC.2019.2958072

Tong SC, Li YM, Sui S, 2016. Adaptive fuzzy tracking control design for SISO uncertain nonstrict feedback non-linear systems. *IEEE Trans Fuzzy Syst*, 24(6):1441-1454.
https://doi.org/10.1109/TFUZZ.2016.2540058

Tong SC, Sun KK, Sui S, 2018. Observer-based adaptive fuzzy decentralized optimal control design for strict-feedback nonlinear large-scale systems. *IEEE Trans Fuzzy Syst*, 26(2):569-584.
https://doi.org/10.1109/TFUZZ.2017.2686373

Wang AJ, Liu WP, Dong T, et al., 2022. DisEHPPC: enabling heterogeneous privacy-preserving consensus-based scheme for economic dispatch in smart grids. *IEEE Trans Cybern*, 52(6):5124-5135.
https://doi.org/10.1109/TCYB.2020.3027572

Wang H, Li M, 2022. Model-free reinforcement learning for fully cooperative consensus problem of nonlinear multiagent systems. *IEEE Trans Neur Netw Learn Syst*, 33(4):1482-1491.
https://doi.org/10.1109/TNNLS.2020.3042508

Wang SB, 2022. Asymptotic tracking control for nonaffine systems with disturbances. *IEEE Trans Circ Syst II Express Briefs*, 69(2):479-483.
https://doi.org/10.1109/TCSII.2021.3080524

Wang YQ, 2019. Privacy-preserving average consensus via state decomposition. *IEEE Trans Autom Contr*, 64(11):4711-4716.
https://doi.org/10.1109/TAC.2019.2902731

Wang YQ, Lu JQ, Zheng WX, et al., 2021. Privacy-preserving consensus for multi-agent systems via node decomposition strategy. *IEEE Trans Circ Syst I Regular Papers*, 68(8):3474-3484.
https://doi.org/10.1109/TCSI.2021.3081372

Wen GX, Li B, 2022. Optimized leader-follower consensus control using reinforcement learning for a class of second-order nonlinear multiagent systems. *IEEE Trans Syst Man Cybern Syst*, 52(9):5546-5555.
https://doi.org/10.1109/TSMC.2021.3130070

Xie ML, Ding DR, Ge XH, et al., 2022. Distributed platooning control of automated vehicles subject to replay attacks based on proportional integral observers. *IEEE/CAA J Autom Sin*, 11(9):1954-1966.
https://doi.org/10.1109/JAS.2022.105941

Xu HY, Ni YH, Liu ZX, et al., 2021. Privacy-preserving leader-following consensus via node-augment mechanism. *IEEE Trans Circ Syst II Express Briefs*, 68(6):2117-2121.
https://doi.org/10.1109/TCSII.2020.3047850

Yan YM, Chen ZY, Varadharajan V, et al., 2021. Distributed consensus-based economic dispatch in power grids using the Paillier cryptosystem. *IEEE Trans Smart Grid*, 12(4):3493-3502.
https://doi.org/10.1109/TSG.2021.3063712

Yang XD, Zhang H, Wang ZP, 2022. Data-based optimal consensus control for multi-agent systems with policy gradient reinforcement learning. *IEEE Trans Neur Netw Learn Syst*, 33(8):3872-3883.
https://doi.org/10.1109/TNNLS.2021.3054685

Yang ZW, Yu LY, Liu YR, et al., 2022. Event-triggered privacy-preserving bipartite consensus for multi-agent systems based on encryption. *Neurocomputing*, 503:162-172. https://doi.org/10.1016/j.neucom.2022.06.074

Yin TJ, Lv YZ, Yu WW, 2020. Accurate privacy preserving average consensus. *IEEE Trans Circ Syst II Express Briefs*, 67(4):690-694.
https://doi.org/10.1109/TCSII.2019.2918709

Yu T, Ma L, Zhang HW, 2019. Prescribed performance for bipartite tracking control of nonlinear multi-agent systems with hysteresis input uncertainties. *IEEE Trans Cybern*, 49(4):1327-1338.
https://doi.org/10.1109/TCYB.2018.2800297

Zhang HG, Jiang H, Luo YH, et al., 2017. Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method. *IEEE Trans Ind Electron*, 64(5):4091-4100. https://doi.org/10.1109/TIE.2016.2542134

Zhang JX, Li KW, Li YM, 2021. Output-feedback based simplified optimized backstepping control for strict-feedback systems with input and state constraints. *IEEE/CAA J Autom Sin*, 8(6):1119-1132.
https://doi.org/10.1109/JAS.2021.1004018

Zhang P, Huang T, Sun XQ, et al., 2023. Privacy-preserving and outsourced multi-party K-means clustering based on multi-key fully homomorphic encryption. *IEEE Trans Dependab Secure Comput*, 20(3):2348-2359.
https://doi.org/10.1109/TDSC.2022.3181667

Zhang XM, Han QL, Ge XH, et al., 2023. Sampled-data control systems with non-uniform sampling: a survey of methods and trends. *Annu Rev Contr*, 55:70-91.
https://doi.org/10.1016/j.arcontrol.2023.03.004

Zhou Q, Li HY, Wang LJ, et al., 2018. Prescribed performance observer-based adaptive fuzzy control for nonstrict-feedback stochastic nonlinear systems. *IEEE Trans Syst Man Cybern Syst*, 48(10):1747-1758. https://doi.org/10.1109/TSMC.2017.2738155

Zuo XJ, Li LX, Peng HP, et al., 2021. Privacy-preserving multidimensional data aggregation scheme without trusted authority in smart grid. *IEEE Syst J*, 15(1):395-406. https://doi.org/10.1109/JSYST.2020.2994363

List of supplementary materials

1 Example 2

Fig. S1 Output consensus tracking performance of $\breve{x}_i$, $\breve{y}_i$, and $\breve{\psi}_i$ in Example 2

Fig. S2 Two-dimensional output consensus tracking in Example 2

Fig. S3 Comparison of follower 2's decryption errors with different methods in Example 2