

# 网络攻击下异构网联系统的分布式自适应动态规划控制

张龙杰, 陈勇<sup>†</sup>, 刘越智, 潘成伟

(电子科技大学 自动化工程学院, 四川 成都 611731; 四川省电动汽车动力系统与安全工程技术研究中心, 四川 成都 611731)

**摘要:** 本文考虑了节点注入攻击下异构网联系统的安全状态估计与控制问题, 通过设计一种基于分布式远程状态安全估计器的有限时间自适应动态规划控制策略, 抑制节点注入攻击对分布式系统协同跟踪效果的影响, 实现对异构网联系统的安全控制. 首先, 为了实现对节点注入攻击下异构网联系统状态信息重塑, 融合最优攻击补偿策略设计, 设计基于无迹卡尔曼滤波的分布式远程状态安全估计器; 然后, 融合远程状态估计器的安全优化目标和协同优化目标, 基于哈密顿方程的最优控制理论, 提出分布式安全优化控制策略; 在此基础上, 基于有限时间优化理论, 提出基于策略迭代算法的有限时间评价-执行网络权重更新算法, 实现对最优控制策略和价值函数的有限时间趋近; 最后, 利用仿真研究和对比分析验证了所提控制策略的有效性.

**关键词:** 最优控制; 网络攻击; 异构网联系统; 分布式控制; 自适应动态规划; 有限时间评价-执行网络算法

**引用格式:** 张龙杰, 陈勇, 刘越智, 等. 网络攻击下异构网联系统的分布式自适应动态规划控制. 控制理论与应用, 2025, 42(4): 669 – 678

DOI: 10.7641/CTA.2023.21026

## Distributed adaptive dynamic programming for heterogeneous interconnected systems under cyber attacks

ZHANG Long-jie, CHEN Yong<sup>†</sup>, LIU Yue-zhi, PAN Cheng-wei

(School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu Sichuan 611731, China; Sichuan Province Engineering Technology Research Center for Electric Vehicle Driving System and Safety, Chengdu Sichuan 611731, China)

**Abstract:** This article considers the secure state estimation and control for the heterogeneous interconnected systems under node-injected attacks, and a finite-time adaptive dynamic programming based on the distributed secure remote estimation is designed to improve the security of the control systems. Firstly, to recover the state information of the heterogeneous interconnected systems under the node-injected attacks, the distributed secure estimator is designed based on the unscented Kalman filter and the optimal attack compensation strategy. Then, by combining the optimal objective of the secure estimation and the optimal objective of the consensus, the distributed secure optimal control strategy is presented based on the solution of the Hamilton equation. Furthermore, to approximate the optimal controller and value function in the finite-time, the finite-time tuning laws of the critic-actor network are proposed based on the policy iteration algorithm and finite-time optimization. Finally, the effectiveness of the proposed method is verified by the comparison result analysis.

**Key words:** optimal control; cyber-attacks; heterogeneous interconnected systems; distributed control; ADP; finite-time critic-actor algorithm

**Citation:** ZHANG Longjie, CHEN Yong, LIU Yuezhi, et al. Distributed adaptive dynamic programming for heterogeneous interconnected systems under cyber attacks. *Control Theory & Applications*, 2025, 42(4): 669 – 678

## 1 引言

近年来, 随着计算机和通信网络技术的发展, 各个工业领域涌现出了以异构网联系统为基础的复杂系

统, 如网联汽车系统、多域电力系统和跨域无人系统等<sup>[1]</sup>. 利用分布式通信网络将各个异构子系统进行互联, 高效地完成复杂的工作目标. 然而, 分布式通信网

收稿日期: 2022-11-21; 录用日期: 2023-12-11.

<sup>†</sup>通信作者. E-mail: ychencd@uestc.edu.cn; Tel.: +86 15982015001.

本文责任编辑: 王龙.

国家重点研发计划项目(2022YFE0120700), 国家自然科学基金项目(61973331, 61973257, 61903064), 四川省科学与技术支持项目基金项目(2021YFG0079, 2021YFG0080, 2021YFG0082)资助.

Supported by the National Key Research and Development Plan Programs of China (2022YFE0120700), the National Nature Science Foundation of China (61973331, 61973257, 61903064) and the Scientific and Technical Supporting Programs of Sichuan Province of China (2021YFG0079, 2021YFG0080, 2021YFG0082).

络的引入,使得异构网联系统对于网络攻击极为敏感<sup>[2]</sup>.因此,如何应对网络攻击和异构特性对网联系统协同性带来的挑战,受到越来越多研究者的关注<sup>[3]</sup>.

异构特性下网联系统子系统结构和参数差异对分布式协同具有不利影响,为了应对这一问题,文献[4-5]提出了在异构线性网联系统下的分布式控制策略.由于具有异构特性的非线性网联系统更为普遍,文献[6]考虑了具有非线性形式的异构网联系统,并提出了一种有向图拓扑下的分布式协同性控制策略.进一步考虑具有未知非线性动态的异构网联系统,Modares等人<sup>[7]</sup>提出了一种优化输出协同控制策略,在未知动态下实现异构网联系统协同控制.然而,开放且异构的复杂网络环境,使得异构网联系统安全威胁日益增加,如何处理网络攻击与异构特性对网联系统协同的叠加影响,对现有方法具有一定挑战.

为了保证异构网联系统的协同安全性,首先,需要对系统状态信息进行估计重塑.文献[8]提出了一种分布式状态融合估计的方式,以应对拒绝服务攻击对系统状态信息可用性的破坏.进一步考虑分布式状态估计的有限时间收敛问题,文献[9]设计了一种有限时间分布式状态安全估计器,融合有限时间反步控制律,提升受攻击系统的安全性.此外,无线通信网络下的远程状态估计具有鲁棒性强、可移动性强和成本低等特点<sup>[10]</sup>,广泛存在于现代控制系统中,Liu等人<sup>[11]</sup>通过设计一种虚假数据注入攻击,分析网络攻击对无线通信网络下分布式远程状态估计器数据可信性的影响.然而,现有文献针对远程状态估计器的工作,主要集中于设计优化隐蔽攻击,且主要以线性远程状态估计器为研究对象,以探索隐蔽特性下网络攻击对远程状态估计器产生的最大影响<sup>[12-13]</sup>.而异构、非线性和网络攻击多重影响导致网联系统状态的信息重塑困难,对现有远程状态估计方法具有严峻的挑战.因此,需要进一步考虑非线性特性、分布式特性和异构特性的远程状态安全估计策略研究.

进一步考虑协同优化控制策略的设计,强化学习和自适应动态规划分布式最优控制策略,已经成为目前研究的热点之一<sup>[14-15]</sup>.如文献[16]针对大规划网联系统的干扰和饱和问题,提出了一种 $L_2$ 稳定意义下的分布式自适应动态规划控制策略.此外,针对具有执行器异构特性的网联系统,Peng等人<sup>[17]</sup>提出了一种基于Q学习的分布式强化学习协同控制策略.进一步考虑到网联系统网络安全性问题,Song等人<sup>[18]</sup>提出了一种基于自适应动态规划的分散式安全控制策略.文献[19]针对执行器和传感器攻击同时存在的情况,提出了一种基于强化学习的安全优化控制框架.分布式结构的通信已广泛应用于网联系统中.文献[20]提出了一种在远程状态估计器下的分布式强化学习安全控制策略,利用估计的状态信息实现对网联系统的安全优化控制.然而,在网络攻击、非线性和异构特性的存

在下,会使得优化控制策略下异构网联系统稳定裕度降低,且渐进收敛的评价-执行网络趋近律,可能导致系统在趋近网络下的不稳定.因此,通过有限时间趋近设计保证趋近网络收敛速度显得尤为关键.

受前述工作的启发,本文利用设计复杂网联系统的优化安全控制策略.本文的主要工作包括:

1) 不同于现有的基于远程状态估计器攻击设计研究,本文综合考虑网络攻击、非线性、异构特性对网联系统的影响,基于分布式无迹卡尔曼滤波器,设计了一种分布式异构非线性远程状态安全估计器;

2) 进一步地,融合估计误差协方差优化,求解最优估计误差下的攻击补偿策略,设计网络攻击对状态估计冗余项的优化补偿策略;

3) 为了减弱网络攻击下分布式最优控制策略稳定裕度降低对评价-执行网络趋近律影响,在现有渐进收敛意义下的趋近律基础上,融合有限时间优化理论,设计了一种有限时间收敛的评价-执行网络在线学习算法,并从理论上证明所提算法的有限时间收敛性.

## 2 系统与问题描述

### 2.1 预备知识

考虑网联系统的网联通信拓扑结构表示为 $\mathcal{G} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ ,其中 $\mathcal{V}$ 表示节点集,有向图的边集表示为 $\mathcal{W} = \{\varepsilon_{ij} | \{\varepsilon_{ij} = (v_i, v_j)\} \in \mathcal{V} \times \mathcal{V} (i \neq j)\}$ ,其中边的方向指向为 $v_i \leftarrow v_j$ , $\mathcal{E} = [e_{ij}]$ 为邻接矩阵,当 $(v_i, v_j) \in \mathcal{W}$ 时,有 $e_{ij} > 0$ ,否则 $e_{ij} = 0$ .此外,节点 $v_i$ 的所有邻居节点集可表示为 $\mathcal{G}_i = \{v_j \in \mathcal{V} : (v_i, v_j) \in \mathcal{W}\}$ ,特别地, $e_{ii} = 0, \forall i = 1, 2, \dots, N$ .此外,对角度矩阵可表示为 $\mathcal{D} = \text{diag}\{d_1, d_2, \dots, d_N\}$ ,其中 $d_i = \sum_{j \in \mathcal{G}_i} e_{ij}$ ,从而可以得到拉普拉斯矩阵 $\mathcal{L} = \mathcal{D} - \mathcal{E}$ . $\text{Tr}(\cdot)$ 表示矩阵的迹.

### 2.2 节点注入攻击下的异构网联系统模型

考虑具有如下非线性动力学方程复杂网联系统:

$$\begin{cases} x_{i,k+1} = f_i(x_{i,k}) + g_i(x_{i,k})u_{i,k} + v_{i,k}, \\ y_{i,k} = C_{i,k}x_{i,k} + \omega_{i,k}, \end{cases} \quad (1)$$

其中: $y_{i,k} \in \mathbb{R}^\phi$ , $x_{i,k} \in \mathbb{R}^n$ , $u_{i,k} \in \mathbb{R}^m$ 分别系统传感器输出、状态和控制输入; $f_i: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , $g_i: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ 表示连续非线性函数; $C_{i,k} \in \mathbb{R}^{\phi \times n}$ 表示传感器输出矩阵.系统干扰和测量噪声满足零均值的高斯分布 $\mathbf{E}(v_{i,k}v_{i,k}^T) = Q_{i,k}$ , $\mathbf{E}(\omega_{i,k}\omega_{i,k}^T) = R_{i,k}$ .

本文仅考虑系统节点的攻击.如文献[21],令 $\mathcal{A}^N(x_{i,k}) \in \mathbb{R}^m$ 表示传感器网络未知有界节点注入攻击函数,则相应攻击模型为

$$u_{i,k}^a = u_{i,k} + \mathcal{A}^N(x_{i,k}), \quad (2)$$

令 $\mathcal{A}^s(x_{i,k}) \in \mathbb{R}^\phi$ 表示传感器网络节点注入攻击函数

$$y_{i,k}^a = y_{i,k} + \mathcal{A}^s(x_{i,k}), \quad (3)$$

令  $\alpha_{i,k+1}^a = g_i(x_{i,k})\mathcal{A}^s(x_{i,k})$  表示动态方程的节点注入攻击引起的冗余项. 为了实现网联系统的协同性, 考虑具有如下形式的复杂网联系统领导者节点模型:

$$x_{0,k+1} = f_0(x_{0,k}), \quad (4)$$

进而定义网联子系统的跟踪误差为  $\zeta_{i,k} = x_{i,k} - x_{0,k}$ , 联立方程(1)(4), 从而可得跟踪误差动态方程为

$$\begin{aligned} \zeta_{i,k+1} = & f_i(x_{i,k}) - f_i(x_{0,k}) + \alpha_{i,k} + \\ & g_i(x_{i,k})u_{i,k} + v_{i,k}, \end{aligned} \quad (5)$$

其中  $\alpha_{i,k} = f_i(x_{0,k}) - f_0(x_{0,k}) + g_i(x_{i,k})\mathcal{A}^N(x_{i,k})$  表示由于节点注入攻击和异构特性引起的系统偏置项. 首先, 给出复杂网联系统的协同性误差为

$$\xi_{i,k} = \sum_{j \in \mathcal{G}_i} e_{ij}(x_{i,k} - x_{j,k}) + \bar{b}_i(x_{i,k} - x_{0,k}), \quad (6)$$

其中:  $\bar{b}_i$  表示牵引节点增益, 且满足  $\bar{b}_i = \begin{cases} b, & i \in \mathcal{V}_u, \\ 0, & i \in \mathcal{V}_M, \end{cases}$   $b$  为一个正数,  $\mathcal{V}_u$  表示不匹配节点集,  $\mathcal{V}_M$  表示匹配节点集. 利用跟踪误差(5)表示式(6)中的协同性误差, 即

$$\xi_{i,k} = (\bar{d}_i + \bar{b}_i)\zeta_{i,k} - \sum_{j \in \mathcal{G}_i} e_{ij}\zeta_{j,k}, \quad (7)$$

其中  $\bar{d}_i = \sum_{j \in \mathcal{G}_i} e_{ij}$ . 通过分析跟踪误差方程(5)以及协同性误差(7), 可以得到整体系统协同性误差方程

$$\xi_k = ((\mathcal{L} + \mathcal{B}) \otimes I_n)\zeta_k, \quad (8)$$

其中  $\zeta_k = x_k - x_{0,k}$  表示全局参考轨迹跟踪误差, 则  $\xi_k = ((\mathcal{L} + \mathcal{B}) \otimes I_n)(x_k - x_{0,k})$ , 其中  $x_{0,k} = \mathbf{1}_n \otimes x_{0,k}$ , 且有  $x_k = [x_{1,k}^T \ x_{2,k}^T \ \cdots \ x_{2,k}^T]^T$ ,  $\mathcal{B} = \text{diag}\{\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_i, \cdots\}$ . 因此, 通过设置牵引权重, 令  $\mathcal{B} + \mathcal{L}$  为可逆矩阵, 则可以优化协同误差(7), 实现网联系统子系统对领导者节点(4)的跟踪.

### 2.3 分布式远程状态估计器

针对如式(1)的非线性动态方程复杂网联系统, 采用基于分布式无迹卡尔曼滤波的远程状态安全估计器. 定义  $\hat{P}_{i,k+1|k}$ ,  $\hat{P}_{i,k+1|k+1}$ ,  $\hat{P}_{i,xy,k+1}$ ,  $\hat{P}_{i,yy,k+1}$  分别为先验状态估计协方差矩阵  $P_{i,k+1|k} = [(x_{i,k+1} - \hat{x}_{i,k+1|k})(x_{i,k+1} - \hat{x}_{i,k+1|k})^T]$ 、后验状态误差协方差矩阵  $P_{i,k+1|k+1} = [(x_{i,k+1} - \hat{x}_{i,k+1|k+1})(x_{i,k+1} - \hat{x}_{i,k+1|k+1})^T]$ 、交叉误差矩阵  $P_{i,xy,k+1} = [(x_{i,k+1} - \hat{x}_{i,k+1|k})(y_{i,k+1} - \hat{y}_{i,k+1|k})^T]$ 、输出误差矩阵  $P_{i,yy,k+1} = [(y_{i,k+1} - \hat{y}_{i,k+1|k})(y_{i,k+1} - \hat{y}_{i,k+1|k})^T]$  的估计值. 在文献[11, 22]基础上, 设计分布式远程估计器

$$\begin{aligned} \hat{x}_{i,k+1|k+1} = & \hat{x}_{i,k+1|k} + K_{i,k+1} [\hat{y}_{i,k+1} - \hat{y}_{i,k+1|k}] + \\ & H_{i,k+1} \sum_{j \in \mathcal{G}_i} e_{ij} [(\hat{x}_{j,k+1|k} - \hat{x}_{i,k+1|k})] + \hat{\alpha}_{i,k}^a, \end{aligned} \quad (9)$$

其中:  $\hat{x}_{i,k+1|k}$ ,  $\hat{x}_{i,k+1|k+1}$  分别为先验和后验状态估计

值,  $\hat{y}_{i,k+1} = y_{i,k+1}^a - \gamma_{i,k+1}^s$  表示攻击补偿后的测量值,  $\hat{\alpha}_{i,k+1}^a$  为式(5)中系统偏置项的补偿值. 式(9)的状态估计器最优估计增益为

$$\begin{cases} K_{i,k+1} = \hat{P}_{i,xy,k+1} [\hat{P}_{i,yy,k+1}]^{-1}, \\ H_{i,k+1} = \lambda_i \hat{P}_{i,k+1|k}, \end{cases} \quad (10)$$

其中:  $\lambda_i = \frac{\varepsilon}{1 + \text{tr}(\hat{P}_{i,k+1|k}^T \hat{P}_{i,k+1|k})}$ ,  $\varepsilon > 0$  为足够小常数.

为了趋近上述协方差矩阵, 建立如下 sigma 向量:

$$\begin{cases} \chi_{i,k|k}^l = \hat{x}_{i,k|k}, & l = 0, \\ \chi_{i,k|k}^l = \hat{x}_{i,k|k} + \\ \quad (\sqrt{(n+a)\hat{P}_{i,k|k}})^l, & l = 1, 2, \dots, n, \\ \chi_{i,k|k}^l = \hat{x}_{i,k|k} - (\sqrt{(n+a)\hat{P}_{i,k|k}})^l, & l = n+1, n+2, \dots, 2n. \end{cases} \quad (11)$$

根据系统方程(1), 得到 sigma 点的动态更新过程

$$\begin{cases} \chi_{i,k+1|k}^l = f_i(\chi_{i,k|k}^l) + g_i(\chi_{i,k|k}^l)u_{i,k}, \\ \hat{x}_{i,k+1|k} = \sum_{l=0}^{2n} w_l^m \chi_{i,k+1|k}^l. \end{cases} \quad (12)$$

然后融合上述 sigma 点和误差协方差矩阵的定义, 从而可以获取误差协方差矩阵的估计为

$$\begin{aligned} \hat{P}_{i,k+1|k} = & \sum_{l=0}^{2n} [w_l^c (\chi_{i,k+1|k}^l - \hat{x}_{i,k+1|k}) \times \\ & (\chi_{i,k+1|k}^l - \hat{x}_{i,k+1|k})^T] + Q_k. \end{aligned} \quad (13)$$

类似地, 针对测量输出的 sigma 点估计值可以表示为

$$\begin{cases} y_{i,k+1|k}^l = C_{i,k} \chi_{i,k+1|k}^l, \\ \hat{y}_{i,k+1|k} = \sum_{l=0}^{2n} w_l^m y_{i,k+1|k}^l, \end{cases} \quad (14)$$

则输出协方差矩阵和交叉协方差矩阵估计值满足

$$\begin{aligned} \hat{P}_{i,yy,k+1} = & \sum_{l=0}^{2n} [w_l^c (y_{i,k+1|k}^l - \hat{y}_{i,k+1|k}) \times \\ & (y_{i,k+1|k}^l - \hat{y}_{i,k+1|k})^T] + R_{k+1}, \end{aligned} \quad (15)$$

$$\begin{aligned} \hat{P}_{i,xy,k+1} = & \sum_{l=0}^{2n} [w_l^c (\chi_{i,k+1|k}^l - \hat{x}_{i,k+1|k}) \times \\ & (y_{i,k+1|k}^l - \hat{y}_{i,k+1|k})^T], \end{aligned} \quad (16)$$

其中上述估计的权重值可以选取为  $w_0^m = \frac{a}{(n+a)}$ ,  $a = \alpha_e^2(n + k_e) - n$ ,  $w_l^m = w_l^c = \frac{1}{2}(n+a)$ ,  $l = 1, 2, \dots, 2n$ ,  $w_0^c = \frac{a}{[(n+a) + (1 - \alpha_e^2 - \beta_e)]}$ .

设计后验状态误差协方差矩阵迭代更新过程为

$$\hat{P}_{i,k+1|k+1} = \hat{P}_{i,k+1|k} - K_{i,k+1} \hat{P}_{i,xy,k+1}^T, \quad (17)$$

针对式(9)中的攻击补偿项  $\gamma_{i,k+1}^s$  和  $\hat{\alpha}_{i,k+1}^a$ , 需要进一步考虑其优化补偿策略的设计.

**注 1** 上述的分布式远程状态估计器, 为了解决网络

攻击对状态估计的影响,式(7)中包含了待设计的网络攻击冗余估计项,从而实现对抗攻击影响的补偿,相关优化设计过程将在第3节中进行体现。

### 3 分布式有限时间自适应动态规划

#### 3.1 分布式最优控制设计

如式(8),令牵引增益为非零,使得 $(\mathcal{L} + \mathcal{B}) \otimes I_n$ 为满秩矩阵.因此,设计分布式优化目标为

$$J_i(\xi_{i,0}, \alpha_{i,0}) = \sum_{k=0}^{\infty} [L_i(\xi_{i,k}) + U_{ii}(u_{i,k}) + \sum_{j \in \mathcal{G}_i} U_{ij}(u_{j,k})], \quad (18)$$

其中:  $L_i(\xi_{i,k}) = \xi_{i,k}^T \mathcal{W}_i \xi_{i,k}$ ,  $U_{ij}(u_{j,k}) = u_{j,k}^T \mathcal{R}_{ij} u_{j,k}$ ,  $U_{ii}(u_{i,k}) = u_{i,k}^T \mathcal{R}_{ii} u_{i,k}$ . 考虑到状态估计误差,设计如下状态估计优化目标:

$$J_i^a(\mathcal{A}_{i,0}^s) = \sum_{q=0}^{\infty} \beta^q \text{Tr}(\mathcal{Y}_{i,q}), \quad (19)$$

其中  $\mathcal{Y}_{i,k} = \begin{bmatrix} \hat{P}_{i,k|k} & \hat{P}_{i,xy,k} \\ \hat{P}_{i,xy,k}^T & \hat{P}_{i,yy,k+1} \end{bmatrix}$ . 联立上述两式,可得

$$\mathcal{J}_i(\xi_{i,k}, \alpha_{i,k}, \mathcal{A}_{i,k}^s) = J_i(\xi_{i,k}, \alpha_{i,k}) + J_i^a(\mathcal{A}_{i,k}^s), \quad (20)$$

其中  $\beta > 0$  表示优化权重. 令  $\ell_{i,k+1}^x = x_{i,k+1} - \hat{x}_{i,k+1|k+1}$ ,  $\Omega_{i,k} = [\xi_{i,k}^T \ell_{i,k}^T]^T$ . 为了求解上述优化问题,首先,给出值函数

$$V_i(\Omega_{i,k}) = \sum_{q=k}^{\infty} [L_i(\xi_{i,q}) + U_{ii}(u_{i,q}) + \sum_{j \in \mathcal{G}_i} U_{ij}(u_{j,q}) + \beta^{k-q} \text{Tr}(\mathcal{Y}_{i,q})] = V_i^x(\xi_{i,k}) + V_i^a(\ell_{i,k}^x). \quad (21)$$

由于针对远程状态估计器优化的值函数满足

$$\begin{aligned} V_i^a(\ell_{i,k}^x) = & \ell_{i,k}^{xT} \ell_{i,k}^x + (y_{i,k+1} - \hat{y}_{i,k+1|k+1} - \gamma_{i,k+1}^s)^T \times \\ & (y_{i,k+1} - \hat{y}_{i,k+1|k+1} - \gamma_{i,k+1}^s) + \beta V_i^a \times \\ & [x_{i,k+1} - \hat{x}_{i,k+1|k} - K_{i,k+1}(y_{i,k+1} - \hat{y}_{i,k+1|k}) + \\ & K_{i,k+1} \gamma_{i,k+1}^s - \hat{\alpha}_{i,k+1}^a - \\ & H_{i,k+1} \sum_{j \in \mathcal{G}_i} e_{ij}(\hat{x}_{j,k+1|k} - \hat{x}_{i,k+1|k})]. \end{aligned} \quad (22)$$

令补偿估计项为优化变量,求解式(22)最小化问题,从而可得式(9)中的节点注入攻击的最优补偿策略为

$$\gamma_{i,k+1}^s = (y_{i,k+1} - \hat{y}_{i,k+1|k+1}) - \frac{1}{2} \beta K_{i,k+1}^T \times \frac{\partial V_{i,k+1}^a}{\partial \ell_{i,k+1}^x}, \quad (23)$$

$$\begin{aligned} \hat{\alpha}_{i,k+1}^a = & \hat{x}_{i,k+1|k+1} - \hat{x}_{i,k+1|k} + \\ & K_{i,k+1}(\hat{y}_{i,k+1|k} - \hat{y}_{i,k+1|k+1}) - \\ & \frac{1}{2} \beta K_{i,k+1} K_{i,k+1}^T \frac{\partial V_{i,k+1}^a}{\partial \ell_{i,k+1}^x} - \end{aligned}$$

$$H_{i,k+1} \sum_{j \in \mathcal{G}_i} e_{ij}(\hat{x}_{j,k+1|k} - \hat{x}_{i,k+1|k}). \quad (24)$$

由于值函数(21)的贝尔曼方程可以展开为

$$\begin{aligned} V_i(\Omega_{i,k}) = & \xi_{i,k}^T \mathcal{W}_i \xi_{i,k} + u_{i,k}^T \mathcal{R}_{ii} u_{i,k} + \\ & \sum_{j \in \mathcal{G}_i} u_{j,k}^T \mathcal{R}_{ij} u_{j,k} + \text{Tr}(\mathcal{Y}_{i,k}) + V_i(\Omega_{i,k+1}), \end{aligned} \quad (25)$$

以控制输入  $u_{i,k}$  为优化变量,通过对上式进行优化求解可得最优的控制率为

$$\pi_{i,k} = -\frac{1}{2} (\bar{d}_i + \bar{b}_i) \mathcal{R}_{ii}^{-1} g_i^T(x_{i,k}) \frac{\partial V_{i,k+1}}{\partial \xi_{i,k+1}}. \quad (26)$$

然后针对式(25)–(26)的优化求解过程,将在下一节设计相应的策略迭代在线学习过程。

#### 3.2 分布式有限时间在线执行–评估网络算法设计

由于式(25)–(26)的求解过程中包含大量非线性项.因此引入评价网络的最优趋近

$$V_i(\Omega_{i,k}) = \Pi_i^T \phi(\Omega_{i,k}) + \rho_{i,k}, \quad (27)$$

其中:  $\Pi_i^T$ ,  $\phi(\cdot)$  分别表示评价网络权重和激励函数,  $\rho_{i,k}$  为评价网络逼近误差. 式(27)的评价网络定义为

$$\hat{V}_i(\Omega_{i,k}) = \hat{\Pi}_i^T \phi(\Omega_{i,k}), \quad (28)$$

从而可以获取趋近的哈密尔顿方程为

$$\begin{aligned} H_i(\Omega_{i,k}, u_i, u_j, \hat{\Pi}_{i,k}^T) = & \xi_{i,k}^T \mathcal{W}_i \xi_{i,k} + u_{i,k}^T \mathcal{R}_{ii} u_{i,k} + \\ & \sum_{j \in \mathcal{G}_i} u_{j,k}^T \mathcal{R}_{ij} u_{j,k} + \text{Tr}(\mathcal{Y}_{i,k}) + \hat{\Pi}_{i,k}^T \phi(\Omega_{i,k+1}) - \\ & \hat{\Pi}_{i,k}^T \phi(\Omega_{i,k}) = e_H. \end{aligned} \quad (29)$$

定义趋近误差优化目标函数为

$$\varepsilon_H = \frac{1}{2} e_H^T e_H, \quad (30)$$

利用有限时间优化理论<sup>[23]</sup>和文献[24]的正则化方法,评价网络权重的有限时间自适应更新过程设计为

$$\hat{\Pi}_{i,k+1} = \hat{\Pi}_{i,k} - \delta_i^V \frac{\phi(\Omega_{i,k+1})}{(\chi_{k+1}^T \chi_{k+1} + 1)^2} \text{sig}^\lambda(w_{H,k}), \quad (31)$$

其中:  $\delta_i^V$  表示学习律,  $0 < \lambda < 1$ , 且  $w_{H,k+1} = w_{H,k} + e_{H,k}$ ,  $\chi_{i,k+1} = \phi(\Omega_{i,k+1}) - \phi(\Omega_{i,k})$ .

由式(21)可知,针对值函数的逼近可以划分为

$$\begin{aligned} \hat{V}_i(\Omega_{i,k}) = & \hat{V}_i^x(\xi_{i,k}) + \hat{V}_i^a(\ell_{i,k}^x) = \\ & \hat{\Pi}_i^{xT} \phi(\xi_{i,k}) + \hat{\Pi}_i^{aT} \phi(\ell_{i,k}^x), \end{aligned} \quad (32)$$

其中:  $\hat{\Pi}_i^x = (\hat{\Pi}_i)_{1 \sim n}$ ,  $\hat{\Pi}_i^a = (\hat{\Pi}_i)_{n+1 \sim 2n}$ .

类似地,控制律的执行网络最优估计可以定义为

$$\pi_{i,k}^* = \Xi_i^{cT} \phi_c(\Omega_{i,k}) + \nu_{i,k}^c, \quad (33)$$

其中:  $\Xi_i^c$  和  $\phi_c(\cdot)$  分别为执行网络的权重和激励函数,

$\nu_{i,k}^c$  表示网络逼近误差, 则执行网络趋近函数定义为

$$\hat{\pi}_{i,k} = \hat{\Xi}_{i,k}^{cT} \phi_c(\xi_{i,k}), \quad (34)$$

将式(28)代入式(26), 则评价网络定义的控制策略表示为

$$\begin{aligned} \bar{\pi}_{i,k} = & -\frac{1}{2}(\bar{d}_i + \bar{b}_i) \mathcal{R}_{ii}^{-1} g_i^T(x_{i,k}) \left[ \frac{\partial \phi(\xi_{i,k+1})}{\partial \xi_{i,k+1}} \right]^T \hat{\Pi}_{i,k+1}^x. \quad (35) \end{aligned}$$

定义优化执行网络与评价网络趋近误差为

$$\begin{aligned} \tilde{\pi}_{i,k} = & \hat{\Xi}_{i,k}^{cT} \phi_c(\xi_{i,k}) + \frac{1}{2}(\bar{d}_i + \bar{b}_i) \mathcal{R}_{ii}^{-1} g_i^T(x_{i,k}) \times \\ & \left[ \frac{\partial \phi(\xi_{i,k+1})}{\partial \xi_{i,k+1}} \right]^T \hat{\Pi}_{i,k+1}, \quad (36) \end{aligned}$$

其中  $\tilde{\pi}_{i,k} = \hat{\pi}_{i,k} - \bar{\pi}_{i,k}$ , 得到执行网络权重优化目标

$$\ell_H = \frac{1}{2} \tilde{\pi}_{i,k}^T \tilde{\pi}_{i,k}, \quad (37)$$

同样, 设计执行网络权重的有限时间更新律为

$$\begin{aligned} \hat{\Xi}_{i,k+1}^c = & \hat{\Xi}_{i,k}^c - \delta_{i,c} (\phi_c^T(\xi_{i,k}) \phi_c(\xi_{i,k}) + 1)^{-1} \phi_c(\xi_{i,k}) \times \end{aligned}$$

$$\begin{aligned} \text{sig}^\lambda \left[ \frac{1}{2} \lambda \mathcal{R}_{ii}^{-1} g_i^T(x_{i,k}) \left[ \frac{\partial \phi(\xi_{i,k+1})}{\partial \xi_{i,k+1}} \right]^T \hat{\Pi}_{i,k+1} + \right. \\ \left. \hat{\Xi}_{i,k}^{cT} \phi_c(\xi_{i,k}) \right]^T, \quad (38) \end{aligned}$$

其中  $\delta_{i,c}$  为正常数, 表示执行网络更新的学习率. 从而, 可以得到评价网络下的节点攻击补偿

$$\begin{aligned} \gamma_{i,k+1}^s = & (y_{i,k+1} - \hat{y}_{i,k+1|k+1}) - \\ & \frac{1}{2} \beta K_{i,k+1}^T \left[ \frac{\partial \phi(\Omega_{i,k+1})}{\partial \ell_{i,k+1}^x} \right]^T \hat{\Pi}_{i,k+1}^a, \quad (39) \end{aligned}$$

$$\begin{aligned} \hat{\alpha}_{i,k+1}^a = & \hat{x}_{i,k+1|k+1} - \hat{x}_{i,k+1|k} + \\ & K_{i,k+1}(\hat{y}_{i,k+1|k} - \hat{y}_{i,k+1|k+1}) - \\ & \frac{1}{2} \beta K_{i,k+1} K_{i,k+1}^T \left[ \frac{\partial \phi(\Omega_{i,k+1})}{\partial \ell_{i,k+1}^x} \right]^T \hat{\Pi}_{i,k+1}^a - \\ & H_{i,k+1} \sum_{j \in i} e_{ij}^a (\hat{x}_{j,k+1|k} - \hat{x}_{i,k+1|k}). \quad (40) \end{aligned}$$

### 4 算法收敛性分析

如图1所示, 本文设计了基于评价网络(31)和执行网络(38)的有限时间自适应动态规划算法. 为了说明本文所提算法的稳定特性, 给出如下引理.

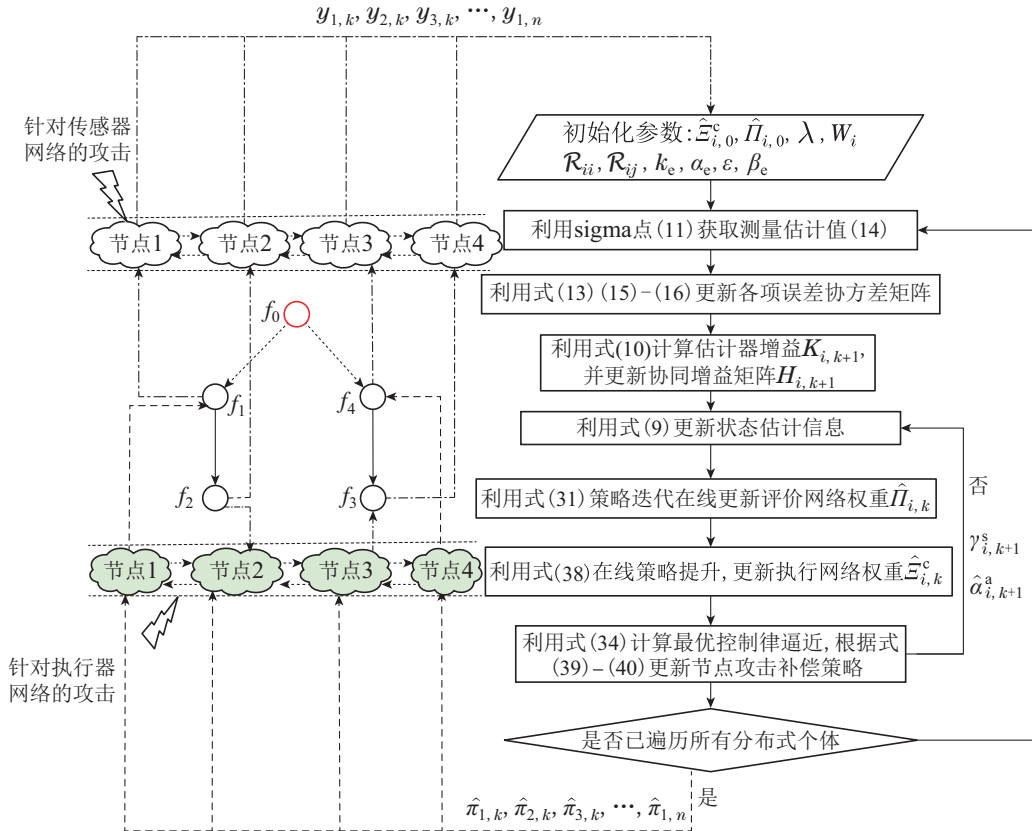


图 1 提出的分布式安全远程状态估计下的有限时间自适应动态规划控制框架

Fig. 1 The proposed finite-time adaptive dynamic programming based on the distributed remote secure estimation

**引理 1** 如文献 [17] 所述, 选取式(21)作为李雅普诺夫函数, 则根据贝尔曼方程 (25) 有  $V_i(\Omega_{i,k+1}) - V_i(\Omega_{i,k}) \leq 0$ , 即在最优控制策略(26)下, 协同误差(7)

为渐进收敛.

给出如下定理说明所提算法对攻击的抑制作用:

**定理 1** 在攻击补偿策略(23)和(24)作用下, 由

攻击引起的状态估计误差函数具有有限上界,即

$$V_i^a(\ell_{i,k}^x) \leq \tilde{V}_{i,0}^{a,0} + \frac{\beta^2 \lambda_{K,\max} \rho_\phi^2}{4(1-\beta)}, \quad (41)$$

其中:  $\lambda_{K,\max}$  表示  $K_{i,k+1} K_{i,k+1}^T$  的最大特征值,  $\rho_\phi$  表示值函数的最大斜率,即满足假设  $\|V(\xi_{i,k}) - V(\xi_{i,0})\|_2 \leq \rho_\phi \|\xi_{i,k} - \xi_{i,0}\|_2$ .

证 将式(23)-(24)代入式(22)中,则有

$$V_i^a(\ell_{i,k}^x) = \ell_{i,k}^{xT} \ell_{i,k}^x + \left\| \frac{1}{2} \beta K_{i,k+1}^T \frac{\partial V_{i,k+1}^a}{\partial \ell_{i,k+1}^x} \right\|^2 + \beta V_{i,k+1}^a. \quad (42)$$

根据值函数定义可得

$$V_i^a(\ell_{i,k}^x) \leq \lim_{L \rightarrow \infty} \sum_{l=0}^L \beta^l (\ell_{i,k+l}^{xT} \ell_{i,k+l}^x) + \frac{\beta^2 \lambda_{K,\max} \rho_\phi^2}{4(1-\beta)} \leq \tilde{V}_{i,0}^{a,0} + \frac{\beta^2 \lambda_{K,\max} \rho_\phi^2}{4(1-\beta)}. \quad (43)$$

证毕.

从而定理1得到了证明,该定理反映了所提补偿策略能够有效减弱网络攻击下估计误差的过大偏移.

为了说明离散形式权重更新策略的有限时间收敛,涉及离散系统的有限时间稳定性,因此给出引理2.

**引理2** 如文献[25]所述,针对离散系统  $x_{k+1} = f(x_k)$ ,存在连续正定函数  $L^V(x_k) > 0, x_k \neq 0, L^V(0) = 0$ ,满足条件

$$L^V(x_{k+1}) - L^V(x_k) \leq -C \min\left\{\frac{L^V(x_k)}{C}, L^{V,\lambda}(x_k)\right\}, x_k \neq 0, \quad (44)$$

其中:  $C > 0$ ,且  $0 < \lambda < 1$ .则存在有限的时间步长  $K(x_0)$ ,令  $x_{K(x_0)} \rightarrow 0$ .有限时间收敛步长满足

$$K(x_0) \geq \log_{1-CL^V(x_0)^{\lambda-1}} \frac{C^{\frac{1-\lambda}{\lambda}}}{L^V(x_0)} + 1. \quad (45)$$

为了说明算法的有限时间收敛性,给出定理2.

**定理2** 在执行网络权重更新过程(31)作用下,执行网络逼近函数(28)能够在有限时间内收敛至评价网络最优值(27),且收敛时间步长满足

$$K^V(L_0^V) \geq \log_{1-C_\delta L_0^{V,(\lambda-1)}} \frac{C_\delta^{\frac{1-\lambda}{\lambda}}}{L_0^V} + 1, \quad (46)$$

其中:  $C_\delta > 0$ ,且  $0 < \lambda < 1$ .

证 首先,给出值函数收敛特性的李雅普诺夫函数

$$L_k^V = |\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k})|, \quad (47)$$

将式(31)代入由式(47)构成的差分方程,可得

$$L_{k+1}^V - L_k^V = \left| -\delta_i^V \frac{1}{(\chi_{k+1}^T \chi_{k+1} + 1)^2} \phi^T(\Omega_{i,k+1}) \phi(\Omega_{i,k+1}) \times \text{sig}^\lambda(w_{H,k}) + \phi^T(\Omega_{i,k+1})(\hat{\Pi}_{i,k} - \Pi_i) - \rho_{i,k+1} \right| - |\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k})|, \quad (48)$$

在零初值下,  $w_{H,k}$  满足

$$w_{H,k+1} = [\hat{V}_i(\Omega_{i,k+1}) - V_i(\Omega_{i,k+1})],$$

且存在

$$L_{k+1}^V - L_k^V = |\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k}) - \delta_i^V \frac{1}{(\chi_{k+1}^T \chi_{k+1} + 1)^2} \phi^T(\Omega_{i,k+1}) \phi(\Omega_{i,k+1}) \times \text{sig}^\lambda[\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k})]| - |\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k})| \leq |L_k^V - C_\delta L_k^{V,\lambda}| - L_k^V, \quad (49)$$

其中  $C_\delta = \delta_i^V \frac{1}{(\chi_{k+1}^T \chi_{k+1} + 1)^2} \phi^T(\Omega_{i,k+1}) \phi(\Omega_{i,k+1})$ .

由绝对值不等式的基本性质可得

1) 当  $L_k^V < C_\delta^{\frac{1-\lambda}{\lambda}}$  时,则式(49)满足

$$L_{k+1}^V - L_k^V \leq C_\delta L_k^{V,\lambda} - 2L_k^V, \quad (50)$$

由于系数  $C_\delta$  的下界满足  $C_\delta > L_k^{V,1-\lambda}$ ,若使上式恒成立,则式(50)需要满足

$$L_{k+1}^V - L_k^V \leq L_k^{V,(1-\lambda)} L_k^{V,\lambda} - 2L_k^V \leq -L_k^V \rightarrow L_{k+1}^V \leq 0 \rightarrow L_{k+1}^V = 0, \quad (51)$$

当存在  $L_k^V < C_\delta^{\frac{1-\lambda}{\lambda}}$  时,在迭代更新率(31)下,评价网络(28)能够收敛至评价网络(27).

2) 当  $L_k^V \geq C_\delta^{\frac{1-\lambda}{\lambda}}$  时,式(49)满足

$$L_{k+1}^V - L_k^V \leq L_k^{V,(1-\lambda)} L_k^{V,\lambda} - 2L_k^V \leq L_{k+1}^V - L_k^V \leq -\min L_k^V - C_\delta L_k^{V,\lambda}, \quad (52)$$

联立式(51)和式(52),可以得到值函数收敛特性李雅普诺夫函数的差分方程满足

$$L_{k+1}^V - L_k^V \leq -C_\delta \min\left\{\frac{L_k^V}{C_\delta}, L_k^{V,\lambda}\right\}, \quad (53)$$

根据引理2可得,评价网络趋近(28)能够在有限时间步长内收敛至值函数的最优估计,且收敛步长满足

$$K^V(L_0^V) \geq \log_{1-C_\delta L_0^{V,(\lambda-1)}} \frac{C_\delta^{\frac{1-\lambda}{\lambda}}}{L_0^V} + 1. \quad (54)$$

因此,定理2的论述得到了证明. 证毕.

**注2** 定理2给出了评价网络值函数的有限时间收敛过程,即当  $k \geq K^V(L_0^V)$  时,存在  $|\hat{V}_i(\Omega_{i,k}) - V_i(\Omega_{i,k})| < \varepsilon, \forall \varepsilon > 0$ .上式等价于  $|\phi^T(\Omega_{i,k})(\hat{\Pi}_{i,k} - \Pi_i) - \rho_{i,k}| < \varepsilon, \forall k >$

$K^V(L_0^V)$ , 则有  $\min_{1, \dots, N} |\phi^T(\Omega_{i,k})|_{1, \dots, N} |(\hat{\Pi}_{i,k} - \Pi_i)| \leq \varepsilon, \forall k > K^V(L_0^V)$ . 从而可得: 当  $k > K^V(L_0^V)$  时,  $|\hat{\Pi}_{i,k} - \Pi_i| < \frac{\varepsilon}{\min_{1, \dots, N} |\phi^T(\Omega_{i,k})|_{1, \dots, N}}$ . 因此, 当有限时间步长满足  $k > K^V(L_0^V)$  时, 评价网络权重满足  $\hat{\Pi}_{i,k} \rightarrow \Pi_i$ , 即评价网络权重在有限时间步长内可以收敛至最优估计权重.

进一步考虑执行网络权重(38)的更新过程, 则可给出定理3, 以说明执行网络的有限时间收敛性特性.

**定理3** 如式(38)的权重更新律, 执行网络(34)能够在有限时间内收敛至控制律最优估计(33), 且收敛步长满足  $K_c = \max\{K^V(L_0^V), K_c(L_{\pi,0})\}$ , 其中  $K_c(L_{\pi,0}) \geq \log_{1-C_\phi} L_{\pi,0}^{\lambda-1} \left(\frac{C_\phi^{1-\lambda}}{L_{\pi,0}}\right) + 1$ .

**证** 首先, 给出执行网络趋近误差李雅普诺夫函数

$$L_{\pi,k} = |\hat{\pi}_{i,k} - \pi_{i,k}^*|, \quad (55)$$

由更新律优化目标(37)可得

$$\tilde{\pi}_{i,k}^* = \pi_{i,k}^* - \tilde{\pi}_{i,k}^* = 0, \quad (56)$$

将式(56)代入式(55)的差分方程, 可得

$$\begin{aligned} \Delta L_{\pi,k} &= \\ &|\hat{\pi}_{i,k+1} - \pi_{i,k+1}^*| - |\hat{\pi}_{i,k} - \pi_{i,k}^*| = \\ &|\phi_c^T \hat{\Xi}_{i,k+1}^c - \tilde{\pi}_{i,k+1}^*| - |\hat{\pi}_{i,k} - \tilde{\pi}_{i,k}^*|, \end{aligned} \quad (57)$$

将式(38)代入式(57)中, 同式(49), 则差分方程满足

$$\begin{aligned} \Delta L_{\pi,k} &= \\ &|\phi_c^T \hat{\Xi}_{i,k}^c - \tilde{\pi}_{i,k}^* - \\ &\delta_{i,c} \phi_c^T \phi_c \text{sig}^\lambda(\hat{\pi}_{i,k} - \tilde{\pi}_{i,k}^*)| - |\hat{\pi}_{i,k} - \tilde{\pi}_{i,k}^*| \leq \\ &-C_\phi \min\left\{\frac{|\hat{\pi}_{i,k} - \tilde{\pi}_{i,k}^*|}{C_\phi} |\hat{\pi}_{i,k} - \tilde{\pi}_{i,k}^*|^\lambda\right\} + \\ &|\tilde{\pi}_{i,k} - \tilde{\pi}_{i,k}^*|, \end{aligned} \quad (58)$$

其中  $C_\phi = \delta_{i,c} \phi_c^T \phi_c$ , 将式(56)代入式(58)则可得

$$\begin{aligned} \Delta L_{\pi,k} &\leq \\ &-C_\phi \min\left\{\frac{L_{\pi,k}}{C_\phi}, L_{\pi,k}^\lambda\right\} + |\tilde{\pi}_{i,k} - \tilde{\pi}_{i,k}^*|, \end{aligned} \quad (59)$$

当时间步长满足  $k > K^V(L_0^V)$  时, 则有  $|\tilde{\pi}_{i,k} - \tilde{\pi}_{i,k}^*| = 0$ . 根据引理2, 式(59)表示了执行网络趋近的收敛时间步长满足  $K_c = \max\{K^V(L_0^V), K_c(L_{\pi,0})\}$ , 其中

$K_c(L_{\pi,0}) \geq \log_{1-C_\phi} L_{\pi,0}^{\lambda-1} \frac{C_\phi^{1-\lambda}}{L_{\pi,0}} + 1$ . 则定理3得到了证明. 证毕.

## 5 仿真与对比分析

考虑具有异构离散时间形式的状态空间方程<sup>[15]</sup>:

$$\begin{cases} x_{i1}(k+1) = ax_{i2}(k) + x_{i1}(k) + v_{i1}(k), \\ x_{i2}(k+1) = -b_1x_{i1}(k) - b_2x_{i1}(k)^3 + \\ \quad x_{i2}(k) + g_iu_i(k) + v_{i2}(k), \\ y_i(k) = [x_{i1}(k) \quad x_{i2}(k)]^T + \omega_i(k), \end{cases} \quad (60)$$

其中系统状态参数满足  $a = 0.05, b_1 = 0.005, b_2 = 0.0035$ , 且  $g_i = [0.06 \quad 0.08 \quad 0.04 \quad 0.02]$ ,  $i = 1, 2, 3, 4, g_0 = 0$ , 分布式系统通信拓扑结构如图1所示. 初始状态可以设置为

$$\begin{aligned} x_1(0) &= \begin{bmatrix} 0.9814 \\ 0.5644 \end{bmatrix}, \quad x_2(0) = \begin{bmatrix} 0.9835 \\ 0.3041 \end{bmatrix}, \\ x_3(0) &= \begin{bmatrix} 0.8458 \\ 0.5084 \end{bmatrix}, \quad x_4(0) = \begin{bmatrix} 0.9824 \\ 0.3317 \end{bmatrix}, \\ x_0(0) &= \begin{bmatrix} 0.8199 \\ 0.4445 \end{bmatrix}. \end{aligned}$$

系统(58)中的干扰和测量噪声满足零均值正态分布, 即  $v_i(k) \sim N(0, Q_i), \omega_i(k) \sim N(0, R_i)$ , 其中  $Q_i = R_i = \begin{bmatrix} 0.001^2 & 0 \\ 0 & 0.001^2 \end{bmatrix}$ . 远程状态估计器的参数为  $\alpha_e = 0.5, k_e = 1, \beta_e = 2, \varepsilon = 0.0001$ . 攻击函数设置为

$$\begin{cases} \mathcal{A}^N(x_{i,k}) = 0.05f_a^A(x_i(k)) + 0.03, \\ \mathcal{A}^S(x_{i,k}) = 0.01f_a^S(y_i(k)) + 0.03. \end{cases}$$

本文仿真实验研究所采用的系统模型拓扑结构和算法流程如图1所示, 攻击施加的采样时刻点为  $k = 350$  和  $k = 750$ . 本文提出的算法是在文献[24]中自适应动态规划算法基础上进行有限时间收敛意义下的改进算法, 因此, 以文献[24]中算法作为参考对比方法.

如图2所示, 在两次攻击发生后, 在所提控制策略作用下, 系统状态  $x_{i1}$  仍然能够在一定采样时刻点内实现对参考轨迹的跟踪; 而在对比方法作用下, 系统状态  $x_{i1}$  在第1次攻击发生后引起的系统状态偏移一直持续到仿真时刻结束.

图3展示了系统状态  $x_{i2}$  的响应情况, 所提方法在两次攻击作用后都能实现对参考轨迹的跟踪, 而所对比方法在第2次攻击作用后, 系统状态对参考轨迹的跟踪则发生了偏移. 上述状态响应情况可以更为直观地反映在图4-5的系统状态轨迹跟踪误差图中, 在本文所提方法下, 两个系统状态在第1次攻击下实现对参考轨迹跟踪的调节时间分别为171个采样时间和187个采样时间; 第2次攻击后的调节时间分别为146个采样时间和139个采样时间.

如图6所示, 在两次攻击作用下优化目标值函数都能在100个采样时刻内收敛至稳态, 展示了所提算法的有限时间收敛特性.

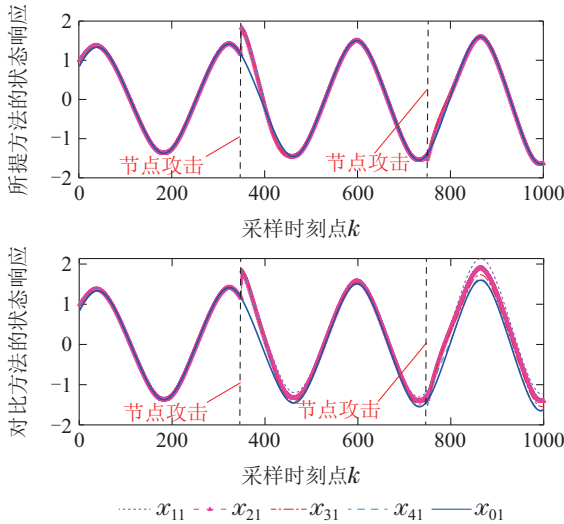


图2 本文所提方法与对比方法下的系统状态 $x_{i1}$ 的响应<sup>[24]</sup>  
 Fig. 2 The states responses of  $x_{i1}$  with proposed method and compared method [24]

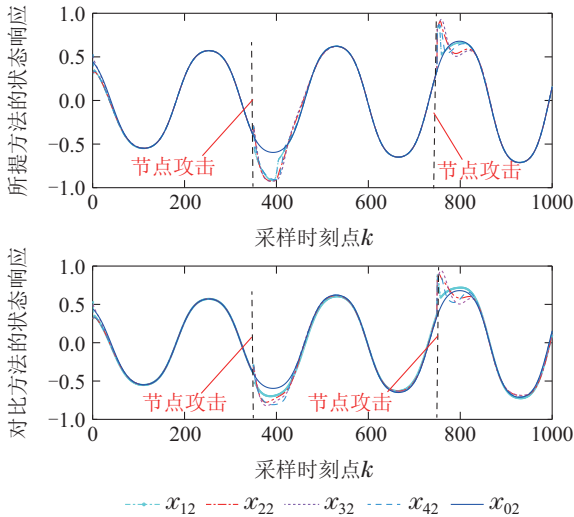


图3 本文所提方法与对比方法下的系统状态 $x_{i2}$ 的响应<sup>[24]</sup>  
 Fig. 3 The states responses of  $x_{i2}$  with proposed method and compared method [24]

图7展示所提控制策略控制输入值情况,其动态过程的调节时间与图4-5展示的协同跟踪误差调节情况同步,显示了所提算法的有效性.图8-9分别展示了所提有限时间自适应动态规划下的评价网络权重和执行网络权重的迭代更新过程.

图10-11所表示的曲线反映了所提远程状态估计器的估计效果,可以看出所有估计状态误差达到一致性,并且两组估计误差上界分别为0.03和0.05,图12展示了先验误差优化更新过程(17),最终达到一致性稳态的效果,体现了所提安全估计器的有效性.

综上所述,在异构网联系统遭受网络攻击时,在本文所提出的分布式安全远程状态估计器下的有限时间自适应动态规划作用下,有效提升分布式异构网联系统的协同轨迹跟踪效果,包括收敛时间和稳态性能方面的提升,通过上述分析展示了所提算法的有效性.

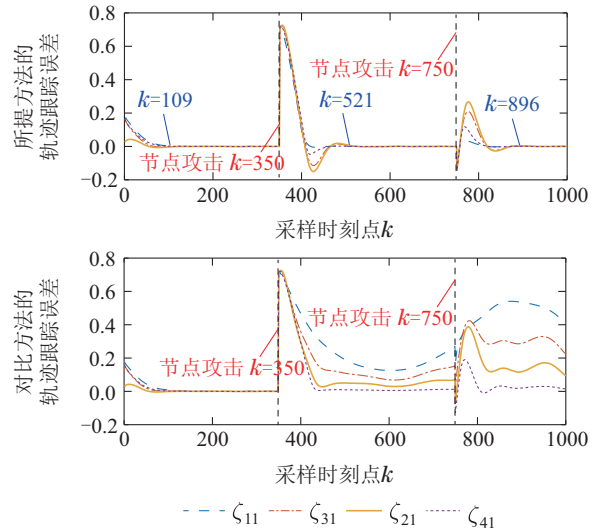


图4 本文所提方法与对比方法下的协同轨迹跟踪误差 $\zeta_{i1}$ <sup>[24]</sup>  
 Fig. 4 The trajectory tracking errors  $\zeta_{i1}$  with the proposed method and compared method [24]

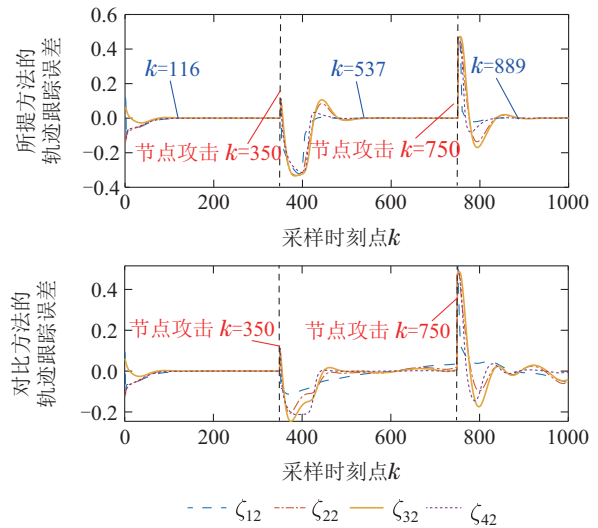


图5 本文所提方法与对比方法下的协同轨迹跟踪误差 $\zeta_{i2}$   
 Fig. 5 The trajectory tracking errors  $\zeta_{i2}$  with the proposed method and compared method

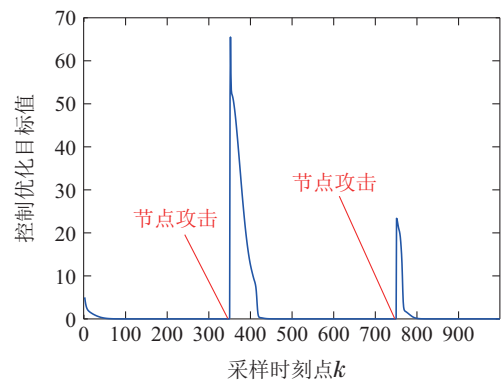


图6 所提方法下的控制策略值函数估计曲线  
 Fig. 6 The critic neural network approximation for the value function with proposed method



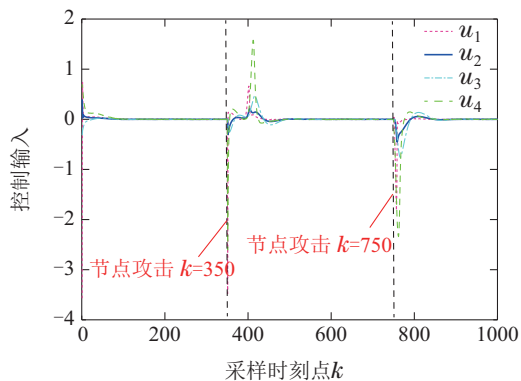


图 7 所提方法的控制输入

Fig. 7 The control input of the proposed method

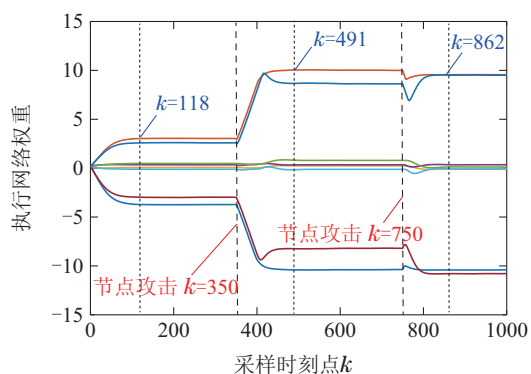


图 8 评价网络权重的在线迭代收敛曲线

Fig. 8 The online iteration process curve of the critic neural network weights

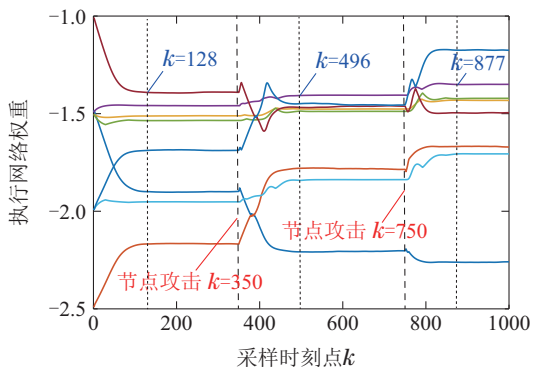


图 9 执行网络权重的在线迭代收敛曲线

Fig. 9 The online iteration process curve of the actor neural network weights

## 6 结论

本文针对节点注入攻击下的异构网联系统, 通过分布式远程状态估计器的设计, 综合设计优化攻击补偿, 从而精确获取无线通信网络下的异构网联系统状态信息, 进而利用有限时间自适应动态规划, 提升异构网联系统状态的分布式协同性能. 仿真结果显示了所提方法的优越性. 虽然本文针对自适应动态规划的执行网络和评价网络的权重更新过程, 设计了一种有限时间收敛算法, 以减弱由于趋近误差导致的系统不

稳定, 然而所设计的控制器仍然是一种渐进稳定控制策略. 此外, 非仿射非线性系统更为广泛地存在于控制系统中<sup>[26]</sup>, 在未来的工作中, 将进一步针对非仿射非线性系统的有限时间优化控制策略的开展研究工作, 以提升优化控制器的动态性能.

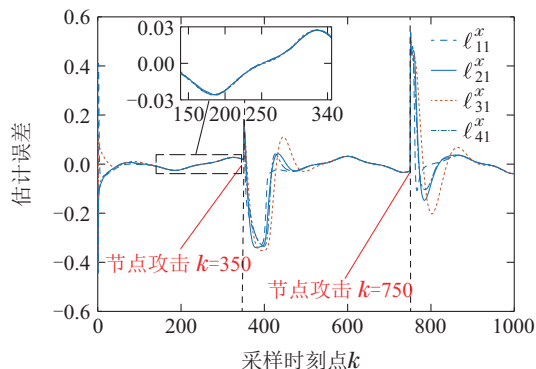


图 10 分布式远程状态估计器估计误差  $\ell_{i1}^x$

Fig. 10 Estimation errors  $\ell_{i1}^x$  of the proposed distributed remote estimator

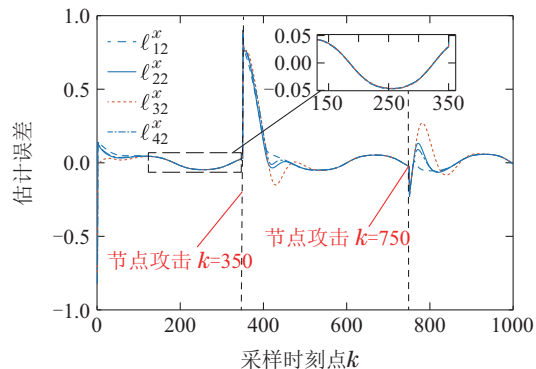


图 11 分布式远程状态估计器估计误差  $\ell_{i2}^x$

Fig. 11 Estimation errors  $\ell_{i2}^x$  of the proposed distributed remote estimator

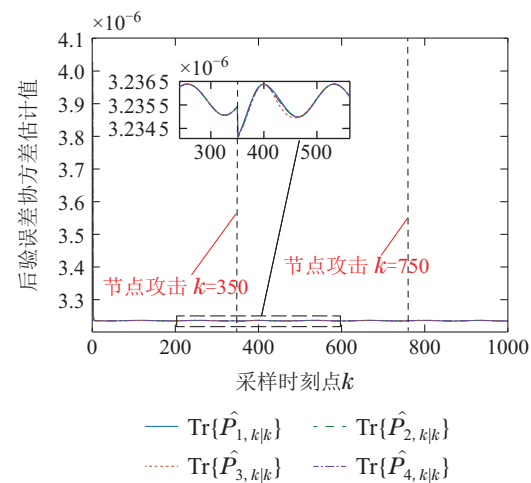


图 12 分布式远程状态估计器先验估计误差协方差的迹

Fig. 12 Trace of covariance of the prior estimation errors with the proposed distributed remote estimator

## 参考文献:

- [1] ZHANG H, LIU C, SU H, et al. Echo state network-based decentralized control of continuous-time nonlinear large-scale interconnected systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(10): 6293 – 6303.
- [2] GU Z, PARK J H, YUE D, et al. Event-triggered security output feedback control for networked interconnected systems subject to cyber-attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(10): 6197 – 6206.
- [3] AN L, YANG G. Decentralized adaptive fuzzy secure control for nonlinear uncertain interconnected systems against intermittent DoS attacks. *IEEE Transactions on Cybernetics*, 2019, 49(3): 827 – 838.
- [4] STÜRZ Y R, EICHLER A, SMITH R S. Distributed control design for heterogeneous interconnected systems. *IEEE Transactions on Automatic Control*, 2021, 66(11): 5112 – 5127.
- [5] HEINKE S, WERNER H. Distributed control of  $\alpha$ -heterogeneous string interconnected systems. *IEEE Control Systems Letters*, 2022, 6: 3122 – 3127.
- [6] PANTELEY E, LORÍA A. Synchronization and dynamic consensus of heterogeneous networked systems. *IEEE Transactions on Automatic Control*, 2017, 62(8): 3758 – 3773.
- [7] MODARES H, LEWIS F L, KANG W, et al. Optimal synchronization of heterogeneous nonlinear systems with unknown dynamics. *IEEE Transactions on Automatic Control*, 2018, 63(1): 117 – 131.
- [8] CHEN B, HO D W C, ZHANG W A, et al. Distributed dimensionality reduction fusion estimation for cyber-physical systems under DoS attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, 49(2): 455 – 468.
- [9] AO Wei, SONG Yongduan, WEN Changyun. Distributed secure state estimation and control for CPSs under sensor attacks – A finite time approach. *Acta Automatica Sinica*, 2019, 45(1): 174 – 184.  
(敖伟, 宋永端, 温长云. 受攻击信息物理系统的分布式安全状态估计与控制——一种有限时间方法. *自动化学报*, 2019, 45(1): 174 – 184.)
- [10] XIAO Li, SUN Zhigang, HU Xiaoya, et al. Remote state estimation based on sensor networks. *Control Theory & Applications*, 2009, 26(7): 763 – 766.  
(肖力, 孙志刚, 胡晓娅, 等. 基于传感器网络的远程状态估计. *控制理论与应用*, 2009, 26(7): 763 – 766.)
- [11] LIU H, NIU B, LI Y. False-data-injection attacks on remote distributed consensus estimation. *IEEE Transactions on Cybernetics*, 2022, 52(1): 433 – 443.
- [12] GUO Z, SHI D, JOHANSSON K, et al. Worst-case stealthy innovation-based linear attack on remote state estimation. *Automatica*, 2018, 89: 117 – 124.
- [13] SHANG J, CHEN M, CHEN T. Optimal linear encryption against stealthy attacks on remote state estimation. *IEEE Transactions on Automatic Control*, 2021, 66(8): 3592 – 3607.
- [14] ZHANG Jilie, ZHANG Huaguang, LUO Yanhong, et al. Nearly optimal control scheme using adaptive dynamic programming based on generalized fuzzy hyperbolic model. *Acta Automatica Sinica*, 2013, 39(2): 142 – 149.  
(张吉烈, 张化光, 罗艳红, 等. 基于广义模糊双曲模型的自适应动态规划最优控制设计. *自动化学报*, 2013, 39(2): 142 – 149.)
- [15] HONG Chengwen, FU Yue. Nonlinear robust approximate optimal tracking control based on adaptive dynamic programming. *Control Theory & Applications*, 2018, 35(9): 1285 – 1292.  
(洪成文, 富月. 基于自适应动态规划的非线性鲁棒近似最优跟踪控制. *控制理论与应用*, 2018, 35(9): 1285 – 1292.)
- [16] TAN L N. Distributed  $H_\infty$  optimal tracking control for strict-feedback nonlinear large-scale systems with disturbances and saturating actuators. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, 50(11): 4719 – 4731.
- [17] PENG Z, LUO R, HU J, et al. Optimal tracking control of nonlinear multiagent systems using internal reinforce Q-learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(8): 4043 – 4055.
- [18] SONG J, HUANG L Y, KARIMI H R, et al. ADP-based security decentralized sliding mode control for partially unknown large-scale systems under injection attacks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2020, 67(12): 5290 – 5301.
- [19] ZHOU Y, VAMVOUDAKIS K G, HADDAD W M, et al. Secure control learning framework for cyber-physical systems under sensor and actuator attacks. *IEEE Transactions on Cybernetics*, 2021, 51(9): 4648 – 4660.
- [20] DAI P, YU W, WANG H, et al. Distributed reinforcement learning for cyber-physical system with multiple remote state estimation under dos attacker. *IEEE Transactions on Network Science and Engineering*, 2020, 7(4): 3212 – 3222.
- [21] JIN X, HADDAD W M, YUCELEN T. An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems. *IEEE Transactions on Automatic Control*, 2017, 62(11): 6058 – 6064.
- [22] OLFATI-SABER R. Distributed Kalman filtering for sensor networks. *The 46th IEEE Conference on Decision and Control*. New Orleans: IEEE, 2007, 12: 5492 – 5498.
- [23] WANG X, WANG G, LI S. Distributed finite-time optimization for integrator chain multiagent systems with disturbances. *IEEE Transactions on Automatic Control*, 2020, 65(12): 5296 – 5311.
- [24] ZHANG H, JIANG H, LUO C, et al. Discrete-time nonzero-sum games for multiplayer using policy-iteration-based adaptive dynamic programming algorithms. *IEEE Transactions on Cybernetics*, 2017, 47(10): 3331 – 3340.
- [25] HADDAD W M, LEE J, BHAT S P. Asymptotic and finite time semistability for nonlinear discrete-time systems with application to network consensus. *IEEE Transactions on Automatic Control*, 2023, 68(2): 766 – 781.
- [26] YANG Hao, PEI Hailong. Nonlinear dynamic inversion approach for a class of nonaffine nonlinear systems with input saturation. *Control Theory & Applications*, 2021, 38(4): 425 – 432.  
(杨浩, 裴海龙. 一类输入饱和和非仿射非线性系统的非线性动态逆控制. *控制理论与应用*, 2021, 38(4): 425 – 432.)

## 作者简介:

**张龙杰** 博士研究生, 目前研究方向为分布式自适应动态规划安全控制、远程状态估计理论, E-mail: lizhang\_uestc@163.com;

**陈勇** 教授, 目前研究方向为跨域无人系统控制与状态估计、工业质量管控知识图谱、过程控制、智能网联汽车安全控制、电机与电力电子控制, E-mail: ychencd@uestc.edu.cn;

**刘越智** 从事博士后研究工作, 目前研究方向为智能微电网分布式控制、事件触发控制理论、安全控制, E-mail: lyz\_zh@163.com;

**潘成伟** 博士研究生, 目前研究方向为智能网联汽车安全控制、固定时间控制理论、预设性能控制理论, E-mail: cwpan\_qy@163.com.