自动化学报
ACTA AUTOMATICA SINICA

非线性多智能体系统事件触发神经网络自适应分布式优化控制

吴畏 李克文 佟绍成

Event-triggered Neural Network Adaptive Distributed Optimal Control for Nonlinear Multi-agent Systems

WU Wei, LI Ke-Wen, TONG Shao-Cheng

在线阅读 View online: https://doi.org/10.16383/j.aas.c250659

非线性多智能体系统事件触发神经网络自适应分布式优化控制

吴 畏 [1] 李克文 [1] 佟绍成 [1]

摘 要 针对具有未知动力学的非线性多智能体系统, 研究事件触发神经网络自适应分布式优化控制问题. 通过结合神经网络与微分图博弈理论, 构建一种新型事件触发神经网络自适应分布式优化控制器. 为解决执行器频繁更新问题, 设计事件触发机制. 建立基于神经网络的强化学习算法, 学习优化控制器与哈密顿–雅可比–贝尔曼方程的解析解, 利用当前采样数据和历史存储数据设计评价网络的权重更新机制. 构造 Lyapunov 函数证明了被控非线性多智能体系统为渐近稳定并达到 Nash 均衡. 计算机仿真结果验证了所提分布式最优控制方案的有效性.

关键词 非线性多智能体系统; 事件触发机制; 分布式神经网络自适应优化控制; Nash 均衡

引用格式 吴畏, 李克文, 佟绍成. 非线性多智能体系统事件触发神经网络自适应分布式优化控制. 自动化学报, 2026, **52**(7): 1334−1346

DOI 10.16383/j.aas.c250659 **CSTR** 32138.14.j.aas.c250659

Event-triggered Neural Network Adaptive Distributed Optimal Control for Nonlinear Multi-agent Systems

WU Wei[1] LI Ke-Wen[1] TONG Shao-Cheng[1]

Abstract This paper investigates the event-triggered neural network (NN) adaptive distributed optimal control problem for nonlinear multi-agent systems with unknown dynamics. By integrating NN with differential graphical game theory, a novel event-triggered NN adaptive distributed optimal controller is developed. To deal with the problem of frequent updates of the actuator, an event-triggered mechanism is designed. A reinforcement learning algorithm based on NN is designed to learn the optimal controller and the analytical solution of the Hamilton-Jacobi-Bellman equation, where the updating mechanisms for the critic network weights are designed by integrating current sampled data and historical stored data. The asymptotic stability of the controlled nonlinear multi-agent systems and the achievement of Nash equilibrium are proved via Lyapunov function. Computer simulation results demonstrate the effectiveness of the proposed distributed optimal control scheme.

Keywords nonlinear multi-agent system; event-triggered mechanism; distributed neural network adaptive optimal control; Nash equilibrium

Citation Wu Wei, Li Ke-Wen, Tong Shao-Cheng. Event-triggered neural network adaptive distributed optimal control for nonlinear multi-agent systems. *Acta Automatica Sinica*, 2026, **52**(7): 1334−1346

在过去十年中, 分布式控制的研究成果取得显著进展. 文献 [1−2] 提出线性多智能体系统的分布式控制方法. 文献 [3−4] 提出非线性多智能体系统的分布式控制策略. 为解决具有未知动态的非线性多智能体系统的分布式控制问题, 文献 [5−7] 通过使用神经网络或模糊逻辑系统来逼近未知函数, 提出神经网络和模糊自适应分布式控制器. 需要指出的是, 尽管上述智能自适应分布式控制方案能够保证被控非线性多智能体系统稳定, 并实现一致性控制目标, 但它们并未考虑优化控制问题.

最优控制不仅要求保证控制系统的稳定性, 还需要最小化控制性能指标. 因此, 非线性多智能体系统的分布式优化控制问题引起控制领域的广泛关注, 并提出多种基于自适应动态规划的智能分布式优化控制设计方法[8−12]. 文献 [8−10] 提出基于神经网络的非线性多智能体系统分布式优化控制方法, 而文献 [11−12] 则提出针对非线性多智能体系统的模糊分布式自适应控制方法. 在这些智能分布式优化控制方案中, 采用执行–评判神经网络 (或模糊) 学习或策略迭代算法来逼近哈密顿–雅可比–贝尔

收稿日期 2025-11-18 录用日期 2026-01-30
Manuscript received November 18, 2025; accepted January 30, 2026
国家自然科学基金 (62503212, 62573216), 辽宁省兴辽英才青年拔尖人才资助计划 (XLYC2503078), 辽宁省科技厅博士启动基金计划项目 (2025-BS-0503), 辽宁省青年科技人才托举工程资助
Supported by National Natural Science Foundation of China (62503212, 62573216), Liaoning Xingliao Talent Plan for Young Top-Notch Talents (XLYC2503078), Liaoning Province Science and Technology Plan Joint Program (2025-BS-0503), and Liaoning Young Elite Scientists Sponsorship Program
本文责任编委 罗彪
Recommended by Associate Editor LUO Biao
1. 辽宁工业大学理学院 锦州 121001
1. College of Science, Liaoning University of Technology, Jinzhou 121001

曼方程的解析解. 然而, 这些智能分布式优化控制方案的不足之处在于, 它们只能最小化局部最优目标, 无法获得每个智能体的本身行为.

微分图博弈为制定战略行为提供合适的解决框架, 其中每个智能体的误差动态和性能指标不仅依赖于自身, 还与其在博弈通信图拓扑中的邻居有关. 在微分图博弈中, 智能体的目标是找到一组可行的控制策略, 即确保系统稳定性的控制策略, 以实现全局同步、全局优化和 Nash 均衡的达成. 基于微分图博弈, 文献 [13−15] 提出线性多智能体系统的分布式优化控制策略, 文献 [16−17] 开发针对非线性多智能体系统的神经网络分布式优化控制策略, 文献 [18−19] 则提出针对非线性多智能体系统的模糊分布式优化控制策略. 这些分布式最优控制策略不仅保证控制的非线性多智能体系统的稳定性, 还实现全局最优性能, 即达到 Nash 均衡.

上述协同优化控制算法均需要频繁更新控制器信号, 这可能导致不必要的数据传输. 事件触发控制因其仅在预设事件阈值发生时才进行信息传输可以解决该问题. 近年来, 文献 [20−22] 针对状态未知的非线性系统, 构建事件触发模糊与神经网络控制问题框架, 所提出的事件触发协同控制方法有效减轻了通信负担并提升了通信资源利用效率. 文献 [23−24] 针对一阶非线性多智能体系统提出事件触发分布式优化控制方法. 虽然文献 [23−24] 研究事件触发分布式优化控制方法, 但它们只能最小化局部最优目标, 无法获取每个智能体的本身行为. 同时, 值得注意的是, 目前尚无关于严格反馈非线性多智能体系统事件触发神经网络自适应分布式优化控制的研究成果, 这也是本研究的另一动机.

基于上述讨论, 本文研究非线性多智能体系统事件触发自适应神经网络分布式优化控制问题. 提出一种事件触发自适应神经网络分布式优化控制框架. 本文的研究贡献如下:

1) 本文首次针对具有未知系统动力学的非线性多智能体系统, 构建事件触发神经网络分布式优化控制框架. 基于该框架, 提出一种事件触发自适应神经网络分布式优化控制方法, 能够确保被控多智能体系统实现渐近稳定并达到 Nash 均衡. 相比之下, 现有微分图博弈非线性多智能体系统的神经网络或模糊分布式优化控制策略[16−19] 均需要频繁更新控制器信号, 而无法在有限的网络通信资源下实现渐近稳定和全局 Nash 均衡.

2) 本文开发用于求解最优控制问题的强化学习算法. 通过采用神经网络作为评价网络, 同步学习最优控制策略与哈密顿−雅可比−贝尔曼方程的解析解. 基于经验回放技术设计评价网络权重的自适应更新律, 所得近似解可在不依赖文献 [16−19] 中“近似误差收敛至零”的假设且无需满足持续激励条件的背景下, 确保被控非线性多智能体系统实现渐近稳定并达到 Nash 均衡.

1　预备知识和问题描述

1.1　网络拓扑图

考虑 M 个智能体构成的多智能体系统, 其通信拓扑描述为图 $\Xi=\{M,\ \bar N\}$, 其中 $M=\{0,\ 1,\ \cdots,\ I\}$ 定义为领导者 (leader) 和 I 个跟随者智能体的集合, $\bar N=\{(l,\ q)\in M\times M\}$ 是节点. 邻接矩阵 $\bar A=[a_{l,\ q}]$, 其中, 如果 $(l,\ q)\in\bar N$, $a_{l,\ q}=1$; 如果 $(l,\ q)\notin\bar N$, $a_{l,\ q}=0$. 图 Ξ 的拉普拉斯矩阵定义为 $\bar L=\Lambda-\bar A$, 其中 $\Lambda=\mathrm{diag}\{\Lambda_1,\ \Lambda_2,\ \cdots,\ \Lambda_M\}$ 和 $\Lambda_l=\sum_{q\in M}a_{l,\ q}$. $\bar E_0=\mathrm{diag}\{\bar E_{1,\ 0},\ \cdots,\ \bar E_{M,\ 0}\}$ 是领导者矩阵, 其中, $\bar E_{l,\ 0}=1$ 表示智能体 l 能够获得领导者信息, 否则, $\bar E_{l,\ 0}=0$.

假设 1[13, 22]**.** 网络拓扑 Ξ 是连通的, 并且至少有一个代理能够获取到领导者.

1.2　问题描述

每个智能体的动力学模型描述为:

$$\begin{cases}\dot x_{l,\ k}=x_{l,\ k+1}+f_{l,\ k}(\bar x_{l,\ k})+d_{l,\ k}(t)\\ \dot x_{l,\ N}=u_l+f_{l,\ N}(\bar x_{l,\ N})+d_{l,\ N}(t)\\ y_l=x_{l,\ 1}\end{cases}\tag{1}$$

其中, $\bar x_{l,\ k}=[x_{l,\ 1},\ x_{l,\ 2},\ \cdots,\ x_{l,\ k}]^{\mathrm T}$ $(l=1,\ 2,\ \cdots,\ M$ 和 $k=1,\ 2,\ \cdots,\ N-1)$ 是第 l 个智能体的状态, N 表示动力学模型的阶数; y_l 和 u_l 是第 l 个智能体的输出和输入; $f_{l,\ k}(\cdot)$ $(f_{l,\ k}(0)=0)$ 是第 l 个智能体的非线性动态并且满足局部 Lipschitz 条件; $d_{l,\ k}(t)$ 为外界扰动并满足 $|d_{l,\ k}(t)|\le d^*_{l,\ k}$, 其中, $d^*_{l,\ k}>0$ 是一个常数.

本文的控制目标是: 对于非线性多智能体系统 (1), 设计基于时变阈值事件触发策略的神经网络自适应混合式协同优化控制律, 使跟随智能体能够渐近跟踪领导智能体的轨迹 y_d, 从而实现编队控制目标, 同时保证一致性跟踪误差渐近收敛到零, 且取得 Nash 均衡控制性能.

2　事件触发神经网络自适应混合式协同优化控制

本节将设计一种事件触发神经网络自适应混合式协同优化控制方法. 该协同控制器由事件触发神

经网络自适应前馈分布式控制器 U_l^a 和事件触发误差反馈神经网络优化协同控制器共同组成. 本文所提出的事件触发神经网络协同方法的框架如图 1 所示.

2.1 事件触发神经元网络自适应分布式前馈控制

本节将通过反步递推控制理论给出事件触发神经网络自适应前馈协同控制.

首先, 定义如下的坐标变换:

$$s_{l,\,1}=\sum_{q=1}^{N}a_{l,\,q}(y_l-y_q)+\bar{E}_{l,\,0}(y_l-y_d) \tag{2}$$

$$s_{l,\,k}=x_{l,\,k}-\alpha_{l,\,k-1} \tag{3}$$

其中, $s_{l,\,1}$ 是协同跟踪误差; $s_{l,\,k}$ 是虚拟误差; $\alpha_{l,\,k-1}$ 为虚拟控制器.

基于坐标变换 (2) ~ (3), 将采用 n 步后步递推控制方法来设计前馈协同控制器. 协同控制器的设计过程如下:

第 1 步. 令 $h_{l,\,1}(s_{l,\,1})=f_{l,\,1}(x_{l,\,1})-f_{l,\,1}(x_{l,\,1d})$; $g_{l,\,1}(x_{l,\,1d})=\sum_{q=1}^{N}(f_{l,\,1}(x_{l,\,1})-f_{q,\,1}(x_{q,\,1})-x_{q,\,2})+\bar{E}_{l,\,0}(f_{l,\,1}(x_{l,\,1})-f_{l,\,1}(x_{l,\,1d}))$; 结合式 (1) ~ (2), 可得:

$$\dot{s}_{l,\,1}=h_{l,\,1}(s_{l,\,1})+g_{l,\,1}(x_{l,\,1d})+s_{l,\,2}+\alpha_{l,\,1}+d_{l,\,1} \tag{4}$$

由于 $g_{l,\,1}(x_{l,\,1d})$ 是一个未知非线性函数, 本文利用 Weierstrass 高阶逼近理论[25], 可采用神经网络 $\hat{g}_{l,\,1}(x_{l,\,1d}|\theta_{l,\,1}^{*})=(\theta_{l,\,1}^{*})^{\mathrm{T}}\varphi_{l,\,1}(x_{l,\,1d})$ 逼近未知非线性函数 $g_{l,\,1}(x_{l,\,1d})$, 并假设

$$g_{l,\,1}(x_{l,\,1d})=(\theta_{l,\,1}^{*})^{\mathrm{T}}\varphi_{l,\,1}(x_{l,\,1d})+\varepsilon_{l,\,1} \tag{5}$$

其中, $\varepsilon_{l,\,1}$ 为最小逼近误差并且满足 $|\varepsilon_{l,\,1}|\le\varepsilon_{l,\,1}^{*}$, $\varepsilon_{l,\,1}^{*}>0$ 是一个未知常数; $\theta_{l,\,1}^{*}$ 为理想权重矩阵;

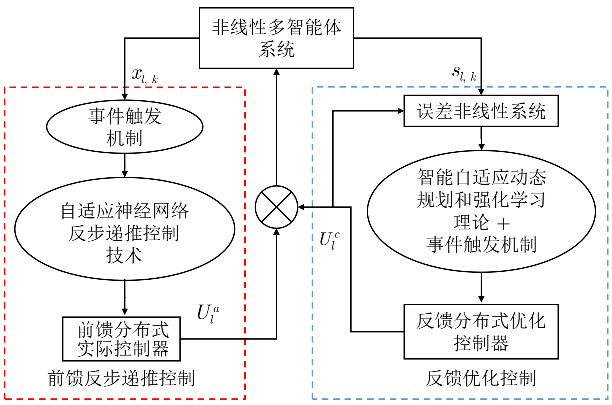

图 1 事件触发分布式优化控制框图

Fig. 1 The block diagram of event-triggered distributed optimal control

$\varphi_{l,\,1}(x_{l,\,1d})$ 为向量激活函数. 需注意函数 $\varphi_l(\cdot)=[\varphi_{l,\,1}(\cdot),\ \varphi_{l,\,2}(\cdot),\ \cdots,\ \varphi_{l,\,N_{f_i}}(\cdot)]^{\mathrm{T}}$ 的选取需满足: 对任意的 $k=1,\ 2,\ \cdots,\ N_{f_i}$, 其分量 $\varphi_{l,\,k}(\cdot)$ 构成一组线性无关的基函数, N_{f_i} 为隐含层节点数.

将式 (5) 代入式 (4), 可得:

$$\begin{aligned}\dot{s}_{l,\,1}=\;&h_{l,\,1}(s_{l,\,1})+(\theta_{l,\,1}^{*})^{\mathrm{T}}\varphi_{l,\,1}(x_{l,\,1d})\,+\\&\varepsilon_{l,\,1}+s_{l,\,2}+x_{l,\,1d}+d_{l,\,1}\end{aligned} \tag{6}$$

构造如下的 Lyapunov 函数

$$V_{l,\,1}=\frac{1}{2}s_{l,\,1}^{2}+\frac{1}{2\sigma_{l,\,1}}\tilde{\Theta}_{l,\,1}^{2}+\frac{1}{2\bar{\sigma}_{l,\,1}}\tilde{\rho}_{l,\,1}^{2} \tag{7}$$

其中, $\tilde{\Theta}_{l,\,1}=\Theta_{l,\,1}^{*}-\hat{\Theta}_{l,\,1}$ 和 $\tilde{\rho}_{l,\,1}=\rho_{l,\,1}^{*}-\hat{\rho}_{l,\,1}$, $\hat{\Theta}_{l,\,1}$ 和 $\hat{\rho}_{l,\,1}$ 是 $\Theta_{l,\,1}^{*}=\|\theta_{l,\,1}^{*}\|$ 和 $\rho_{l,\,1}^{*}=\varepsilon_{l,\,1}^{*}+d_{l,\,1}^{*}$ 的估计; $\bar{\sigma}_{l,\,1}>0$ 和 $\sigma_{l,\,1}>0$ 是常数.

根据式 (2) 和 (3) 并且令 $x_{l,\,1d}=\alpha_{l,\,1}^{a}+\alpha_{l,\,1}^{c}$, $\alpha_{l,\,1}^{a}$ 和 $\alpha_{l,\,1}^{c}$ 分别为前馈反步递推虚拟控制器和优化控制器, 可得:

$$\begin{aligned}\dot{V}_{l,\,1}=\;&s_{l,\,1}\{s_{l,\,2}+h_{l,\,1}(s_{l,\,1})+(\theta_{l,\,1}^{*})^{\mathrm{T}}\varphi_{l,\,1}(x_{l,\,1d})\,+\\&\alpha_{l,\,1}^{c}+\alpha_{l,\,1}^{a}+\rho_{l,\,1}^{*}\}\,+\\&\frac{1}{\sigma_{l,\,1}}\tilde{\Theta}_{l,\,1}\dot{\tilde{\Theta}}_{l,\,1}+\frac{1}{\bar{\sigma}_{l,\,1}}\tilde{\rho}_{l,\,1}\dot{\tilde{\rho}}_{l,\,1}\end{aligned} \tag{8}$$

利用杨氏不等式 $ab\le\frac{a^2}{2}+\frac{b^2}{2}$ (a、$b\in\mathbf{R}$), 可得:

$$\begin{aligned}s_{l,\,1}(\theta_{l,\,1}^{*})^{\mathrm{T}}\varphi_{l,\,1}(x_{l,\,1d})\le\;&\frac{\Theta_{l,\,1}^{*}s_{l,\,1}^{2}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}}{\sqrt{s_{l,\,1}^{2}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}+\eta_{l,\,1}^{2}}}\,+\\&\Theta_{l,\,1}^{*}\eta_{l,\,1}\end{aligned} \tag{9}$$

其中, $\eta_{l,\,1}\in L_2(0,\ \infty)$.

将式 (9) 代入式 (8), 可得:

$$\begin{aligned}\dot{V}_{l,\,1}\le\;&s_{l,\,1}\Bigg\{s_{l,\,2}+\alpha_{l,\,1}^{a}+\alpha_{l,\,1}^{c}+h_{l,\,1}(s_{l,\,1})+\rho_{l,\,1}^{*}\,+\\&\frac{\Theta_{l,\,1}^{*}s_{l,\,1}^{2}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}}{\sqrt{s_{l,\,1}^{2}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}+\eta_{l,\,1}^{2}}}\Bigg\}+\frac{1}{\sigma_{l,\,1}}\tilde{\Theta}_{l,\,1}\dot{\tilde{\Theta}}_{l,\,1}\,+\\&\frac{1}{\bar{\sigma}_{l,\,1}}\tilde{\rho}_{l,\,1}\dot{\tilde{\rho}}_{l,\,1}+\Theta_{l,\,1}^{*}\eta_{l,\,1}\end{aligned} \tag{10}$$

设计如下的虚拟控制器 $\alpha_{l,\,1}^{a}$ 和参数 $\hat{\Theta}_{l,\,1}$、$\hat{\rho}_{l,\,1}$ 的自适应律

$$\begin{aligned}\alpha_{l,\,1}^{a}=\;&-c_{l,\,1}s_{l,\,1}-\frac{\hat{\Theta}_{l,\,1}s_{l,\,1}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}}{\sqrt{s_{l,\,1}^{2}\varphi_{l,\,1}^{\mathrm{T}}\varphi_{l,\,1}+\eta_{l,\,1}^{2}}}\,-\\&\frac{\hat{\rho}_{l,\,1}s_{l,\,1}}{\sqrt{s_{l,\,1}^{2}+\eta_{l,\,1}^{2}}}\end{aligned} \tag{11}$$

$$\dot{\hat{\Theta}}_{l,1}=\sigma_{l,1}\frac{s_{l,1}^2\varphi_{l,1}^{\mathrm{T}}\varphi_{l,1}}{\sqrt{s_{l,1}^2\varphi_{l,1}^{\mathrm{T}}\varphi_{l,1}+\eta_{l,1}^2}}-\sigma_{l,1}\eta_{l,1}\hat{\Theta}_{l,1} \tag{12}$$

$$\dot{\hat{\rho}}_{l,1}=\bar{\sigma}_{l,1}\frac{s_{l,1}^2}{\sqrt{s_{l,1}^2+\eta_{l,1}^2}}-\bar{\sigma}_{l,1}\eta_{l,1}\hat{\rho}_{l,1} \tag{13}$$

式中, $c_{l,1}>0$ 是常数.

将式 (11) ~ (13) 代入式 (10), 可得:

$$\begin{aligned}\dot{V}_{l,1}\le&-c_{l,1}s_{l,1}^2+s_{l,1}\alpha_{l,1}^c+s_{l,1}h_{l,1}(s_{l,1})+\\&\eta_{l,1}\Theta_{l,1}^*+\eta_{l,1}\tilde{\Theta}_{l,1}\hat{\Theta}_{l,1}+\eta_{l,1}\tilde{\rho}_{l,1}\hat{\rho}_{l,1}+\\&|s_{l,1}|\rho_{l,1}^*+s_{l,1}s_{l,2}-\frac{s_{l,1}^2}{\sqrt{s_{l,1}^2+\eta_{l,1}^2}}\end{aligned} \tag{14}$$

利用下列不等式

$$|s_{l,1}|\rho_{l,1}^*-\frac{s_{l,1}^2}{\sqrt{s_{l,1}^2+\eta_{l,1}^2}}\le\eta_{l,1}\rho_{l,1}^* \tag{15}$$

可得:

$$\begin{aligned}\dot{V}_{l,1}\le&-c_{l,1}s_{l,1}^2+s_{l,1}\alpha_{l,1}^c+s_{l,1}h_{l,1}(s_{l,1})+\\&\eta_{l,1}(\rho_{l,1}^*+\Theta_{l,1}^*)+\eta_{l,1}\tilde{\Theta}_{l,1}\hat{\Theta}_{l,1}+\\&\eta_{l,1}\tilde{\rho}_{l,1}\hat{\rho}_{l,1}+s_{l,1}s_{l,2}\end{aligned} \tag{16}$$

第 k 步 ($2\le k\le N-1$)**.** 从式 (3) 可得:

$$\dot{s}_{l,k}=s_{l,k+1}+x_{l,kd}+f_{l,k}(\bar{x}_{l,k})-\dot{x}_{l,(k-1)d}+d_{l,k}(t) \tag{17}$$

令 $h_{l,k}(s_{l,k})=f_{l,k}(\bar{x}_{l,k})-f_{l,k}(x_{l,kd})$ 和 $x_{l,kd}=\alpha_{l,k}^a+\alpha_{l,k}^c$, $\alpha_{l,k}^a$ 和 $\alpha_{l,k}^c$ 分别为前馈反步递推虚拟控制器和优化控制器, 式 (17) 变为如下不等式

$$\begin{aligned}\dot{s}_{l,k}=&s_{l,k+1}+\alpha_{l,k}^a+\alpha_{l,k}^c+h_{l,k}(s_{l,k})+\\&f_{l,k}(x_{l,kd})+d_{l,k}(t)-\dot{x}_{l,(k-1)d}\end{aligned} \tag{18}$$

由于 $F_{l,k}(x_{l,kd})=f_{l,k}(x_{l,kd})-\dot{x}_{l,kd}$ 是一个未知非线性函数, 利用神经元网络 $\hat{F}_{l,k}(x_{l,kd}|\theta_{l,k}^*)=(\theta_{l,k}^*)^{\mathrm{T}}\varphi_{l,k}(x_{l,kd})$ 逼近未知非线性函数 $F_{l,k}(x_{l,kd})$ 并且假设

$$F_{l,k}(x_{l,kd})=(\theta_{l,k}^*)^{\mathrm{T}}\varphi_{l,k}(x_{l,kd})+\varepsilon_{l,k} \tag{19}$$

式中, $\theta_{l,k}^*$ 是理想权重矩阵; $\varepsilon_{l,k}$ 为最小逼近误差并且满足 $|\varepsilon_{l,k}|\le\varepsilon_{l,k}^*$, $\varepsilon_{l,k}^*>0$ 为常数.

令 $\rho_{l,k}^*=\varepsilon_{l,k}^*+d_{l,k}^*$, 并且根据式 (18) 和 (19), 可得:

$$\begin{aligned}\dot{s}_{l,k}=&s_{l,k+1}+\alpha_{l,k}^a+\alpha_{l,k}^c+h_{l,k}(s_{l,k})+\\&(\theta_{l,k}^*)^{\mathrm{T}}\varphi_{l,k}(x_{l,kd})+\rho_{l,k}^*\end{aligned} \tag{20}$$

构造如下的 Lyapunov 函数

$$V_{l,k}=V_{l,k-1}+\frac{1}{2}s_{l,k}^2+\frac{1}{2\sigma_{l,k}}\tilde{\Theta}_{l,k}^2+\frac{1}{2\bar{\sigma}_{l,k}}\tilde{\rho}_{l,k}^2 \tag{21}$$

式中, $\tilde{\Theta}_{l,k}=\Theta_{l,k}^*-\hat{\Theta}_{l,k}$ 和 $\tilde{\rho}_{l,k}=\rho_{l,k}^*-\hat{\rho}_{l,k}$, $\hat{\Theta}_{l,k}$ 和 $\hat{\rho}_{l,k}$ 分别是 $\Theta_{l,k}^*=\|\theta_{l,k}^*\|$ 和 $\rho_{l,k}^*$ 的估计; $\bar{\sigma}_{l,k}>0$ 和 $\sigma_{l,k}>0$ 是常数.

根据式 (20) 和 (21), $V_{l,k}$ 的导数为:

$$\begin{aligned}\dot{V}_{l,k}=&\dot{V}_{l,k-1}+\frac{1}{\sigma_{l,k}}\tilde{\Theta}_{l,k}\dot{\hat{\Theta}}_{l,k}+\frac{1}{\bar{\sigma}_{l,k}}\tilde{\rho}_{l,k}\dot{\hat{\rho}}_{l,k}+\\&s_{l,k}\big[h_{l,k}(s_{l,k})+s_{l,k+1}+\alpha_{l,k}^a+\alpha_{l,k}^c+\\&(\theta_{l,k}^*)^{\mathrm{T}}\varphi_{l,k}(x_{l,kd})+\rho_{l,k}^*\big]\end{aligned} \tag{22}$$

利用杨氏不等式, 可得:

$$\begin{aligned}s_{l,k}(\theta_{l,k}^*)^{\mathrm{T}}\varphi_{l,k}(x_{l,kd})\le&\frac{\Theta_{l,k}^*s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}}{\sqrt{s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}+\eta_{l,k}^2}}+\\&\Theta_{l,k}^*\eta_{l,k}\end{aligned} \tag{23}$$

其中, $\eta_{l,k}\in L_2(0,\infty)$.

根据式 (22) 和 (23), 可得:

$$\begin{aligned}\dot{V}_{l,k}\le&\dot{V}_{l,k-1}+\frac{1}{\bar{\sigma}_{l,k}}\tilde{\rho}_{l,k}\dot{\hat{\rho}}_{l,k}+s_{l,k}\Bigg[h_{l,k}(s_{l,k})+\\&s_{l,k+1}+\alpha_{l,k}^c+\frac{\Theta_{l,k}^*s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}}{\sqrt{s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}+\eta_{l,k}^2}}+\\&\alpha_{l,k}^a+\Theta_{l,k}^*\eta_{l,k}+\rho_{l,k}^*\Bigg]+\frac{1}{\sigma_{l,k}}\tilde{\Theta}_{l,k}\dot{\hat{\Theta}}_{l,k}\end{aligned} \tag{24}$$

设计如下的虚拟控制器 $\alpha_{l,k}^a$ 和参数 $\hat{\Theta}_{l,k}$、$\hat{\rho}_{l,k}$ 的自适应律

$$\begin{aligned}\alpha_{l,k}^a=&-c_{l,k}s_{l,k}-\frac{\hat{\Theta}_{l,k}s_{l,k}\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}}{\sqrt{s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}+\eta_{l,k}^2}}-\\&s_{l,k-1}-\frac{\hat{\rho}_{l,k}s_{l,k}}{\sqrt{s_{l,k}^2+\eta_{l,k}^2}}\end{aligned} \tag{25}$$

$$\dot{\hat{\Theta}}_{l,k}=\sigma_{l,k}\frac{s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}}{\sqrt{s_{l,k}^2\varphi_{l,k}^{\mathrm{T}}\varphi_{l,k}+\eta_{l,k}^2}}-\sigma_{l,k}\eta_{l,k}\hat{\Theta}_{l,k} \tag{26}$$

$$\dot{\hat{\rho}}_{l,k}=\bar{\sigma}_{l,k}\frac{s_{l,k}^2}{\sqrt{s_{l,k}^2+\eta_{l,k}^2}}-\bar{\sigma}_{l,k}\eta_{l,k}\hat{\rho}_{l,k} \tag{27}$$

式中, $c_{l,k}>0$ 是常数.

将式 (25) ~ (27) 代入式 (24), 可得:

$$\dot{V}_{l,\,k} \le -\sum_{j=1}^{k} c_{l,\,j} s_{l,\,j}^2 + \sum_{j=1}^{k} s_{l,\,j}\alpha_{l,\,j}^c + \sum_{j=1}^{k} s_{l,\,j} h_{l,\,j}(s_{l,\,j}) + \sum_{j=1}^{k}\left[\eta_{l,\,j}\Theta_{l,\,j}^* + \eta_{l,\,j}\tilde{\Theta}_{l,\,j}\hat{\Theta}_{l,\,j} + \eta_{l,\,j}\tilde{\rho}_{l,\,j}\hat{\rho}_{l,\,j}\right] + |s_{l,\,k}|\rho_{l,\,k}^* - s_{l,\,1}\sum_{q=1}^{M} a_{l,\,q}\alpha_{q,\,2}^c + s_{l,\,k}s_{l,\,k+1} - \frac{s_{l,\,k}^2}{\sqrt{s_{l,\,k}^2+\eta_{l,\,k}^2}} \tag{28}$$

利用如下不等式

$$|s_{l,\,k}|\rho_{l,\,k}^* - \frac{s_{l,\,k}^2}{\sqrt{s_{l,\,k}^2+\eta_{l,\,k}^2}} \le \eta_{l,\,k}\rho_{l,\,k}^* \tag{29}$$

可得:

$$\dot{V}_{l,\,k} \le -\sum_{j=1}^{k} c_{l,\,j} s_{l,\,j}^2 + \sum_{j=1}^{k} s_{l,\,j}\alpha_{l,\,j}^c + \sum_{j=1}^{k} s_{l,\,j} h_{l,\,j}(s_{l,\,j}) + \sum_{j=1}^{k}\left[\eta_{l,\,j}(\rho_{l,\,j}^* + \Theta_{l,\,j}^*) + \eta_{l,\,j}\tilde{\Theta}_{l,\,j}\hat{\Theta}_{l,\,j} + \eta_{l,\,j}\tilde{\rho}_{l,\,j}\hat{\rho}_{l,\,j}\right] - s_{l,\,1}\sum_{q=1}^{M} a_{l,\,q}\alpha_{q,\,2}^c + s_{l,\,k}s_{l,\,k+1} \tag{30}$$

第 N 步. 根据式 (3) 可得:

$$\dot{s}_{l,\,N} = u_l + f_{l,\,N}(\bar{x}_{l,\,N}) + d_{l,\,N}(t) - \dot{x}_{l,\,(N-1)d} \tag{31}$$

令 $h_{l,\,N}(s_{l,\,N}) = f_{l,\,N}(\bar{x}_{l,\,N}) - f_{l,\,N}(x_{l,\,Nd})$ 和 $u_l = u_l^a + u_l^c$, u_l^a 和 u_l^c 分别为前馈反步递推虚拟控制器和优化控制器, 式 (31) 变为如下不等式

$$\dot{s}_{l,\,N} = u_l^a + u_l^c + h_{l,\,N}(s_{l,\,N}) + f_{l,\,N}(x_{l,\,Nd}) + d_{l,\,N}(t) - \dot{\alpha}_{l,\,N-1} \tag{32}$$

由于 $F_{l,\,N}(\bar{x}_{l,\,Nd}) = f_{l,\,N}(x_{l,\,Nd}) - \dot{x}_{l,\,Nd}$ 是未知非线性函数, 利用神经元网络 $\hat{F}_{l,\,N}(\bar{x}_{l,\,Nd}|\theta_{l,\,N}^*) = (\theta_{l,\,N}^*)^{\rm T}\varphi_{l,\,N}(\bar{x}_{l,\,Nd})$ 逼近未知非线性函数 $F_{l,\,N}(\bar{x}_{l,\,Nd})$, 并假设

$$F_{l,\,N}(\bar{x}_{l,\,Nd}) = (\theta_{l,\,N}^*)^{\rm T}\varphi_{l,\,N}(\bar{x}_{l,\,Nd}) + \varepsilon_{l,\,N} \tag{33}$$

式中, $\theta_{l,\,N}^*$ 是理想权重矩阵; $\varepsilon_{l,\,N}$ 为最小逼近误差并且满足 $|\varepsilon_{l,\,N}| \le \varepsilon_{l,\,N}^*$, $\varepsilon_{l,\,N}^* > 0$ 为常数.

令 $\rho_{l,\,N}^* = \varepsilon_{l,\,N}^* + d_{l,\,N}^*$, 根据式 (32) ~ (33), 可得:

$$\dot{s}_{l,\,N} = u_l^a + u_l^c + h_{l,\,N}(s_{l,\,N}) + (\theta_{l,\,N}^*)^{\rm T}\varphi_{l,\,N}(\bar{x}_{l,\,Nd}) + \rho_{l,\,N}^* \tag{34}$$

构造如下的 Lyapunov 函数

$$V_{l,\,N} = V_{l,\,N-1} + \frac{1}{2}s_{l,\,N}^2 + \frac{1}{2\sigma_{l,\,N}}\tilde{\Theta}_{l,\,N}^2 + \frac{1}{2\bar{\sigma}_{l,\,N}}\tilde{\rho}_{l,\,N}^2 \tag{35}$$

式中, $\tilde{\Theta}_{l,\,N} = \Theta_{l,\,N}^* - \hat{\Theta}_{l,\,N}$ 和 $\tilde{\rho}_{l,\,N} = \rho_{l,\,N}^* - \hat{\rho}_{l,\,N}$, $\hat{\Theta}_{l,\,N}$ 和 $\hat{\rho}_{l,\,N}$ 是 $\Theta_{l,\,N}^* = \|\theta_{l,\,N}^*\|$ 和 $\rho_{l,\,N}^*$ 的估计; $\bar{\sigma}_{l,\,N} > 0$ 和 $\sigma_{l,\,N} > 0$ 是常数.

根据式 (3), $V_{l,\,N}$ 的导数为:

$$\dot{V}_{l,\,N} = \dot{V}_{l,\,N-1} + \frac{1}{\sigma_{l,\,N}}\tilde{\Theta}_{l,\,N}\dot{\hat{\Theta}}_{l,\,N} + \frac{1}{\bar{\sigma}_{l,\,N}}\tilde{\rho}_{l,\,N}\dot{\hat{\rho}}_{l,\,N} + s_{l,\,N}\left[h_{l,\,N}(s_{l,\,N}) + u_l^a + u_l^c + (\theta_{l,\,N}^*)^{\rm T}\varphi_{l,\,N}(\bar{x}_{l,\,Nd}) + \rho_{l,\,N}^*\right] \tag{36}$$

利用杨氏不等式, 可得:

$$s_{l,\,N}(\theta_{l,\,N}^*)^{\rm T}\varphi_{l,\,N}(\bar{x}_{l,\,Nd}) \le \frac{\Theta_{l,\,N}^* s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N}}{\sqrt{s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N} + \eta_{l,\,N}^2}} + \Theta_{l,\,N}^*\eta_{l,\,N} \tag{37}$$

其中, $\eta_{l,\,N} \in L_2(0,\,\infty)$.

将式 (37) 代入式 (36), 可得:

$$\dot{V}_{l,\,N} \le \dot{V}_{l,\,N-1} + \frac{1}{\sigma_{l,\,N}}\tilde{\Theta}_{l,\,N}\dot{\hat{\Theta}}_{l,\,N} + \frac{1}{\bar{\sigma}_{l,\,N}}\tilde{\rho}_{l,\,N}\dot{\hat{\rho}}_{l,\,N} + s_{l,\,N}\left[h_{l,\,N}(s_{l,\,N}) + u_l^a + u_l^c + \frac{\Theta_{l,\,N}^* s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N}}{\sqrt{s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N} + \eta_{l,\,N}^2}} + \Theta_{l,\,N}^*\eta_{l,\,N} + \rho_{l,\,N}^*\right] \tag{38}$$

设计如下的虚拟控制器 $\alpha_{l,\,N}^a$ 和参数 $\hat{\Theta}_{l,\,N}$、$\hat{\rho}_{l,\,N}$ 的自适应律

$$\alpha_{l,\,N}^a = -c_{l,\,N}s_{l,\,N} - \frac{\hat{\Theta}_{l,\,N}s_{l,\,N}\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N}}{\sqrt{s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N} + \eta_{l,\,N}^2}} - s_{l,\,N-1} - \frac{\hat{\rho}_{l,\,N}s_{l,\,N}}{\sqrt{s_{l,\,N}^2 + \eta_{l,\,N}^2}} \tag{39}$$

$$\dot{\hat{\Theta}}_{l,\,N} = \sigma_{l,\,N}\frac{s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N}}{\sqrt{s_{l,\,N}^2\varphi_{l,\,N}^{\rm T}\varphi_{l,\,N} + \eta_{l,\,N}^2}} - \sigma_{l,\,N}\eta_{l,\,N}\hat{\Theta}_{l,\,N} \tag{40}$$

$$\dot{\hat{\rho}}_{l,\,N} = \bar{\sigma}_{l,\,N}\frac{s_{l,\,N}^2}{\sqrt{s_{l,\,N}^2+\eta_{l,\,N}^2}} - \bar{\sigma}_{l,\,N}\eta_{l,\,N}\hat{\rho}_{l,\,N} \tag{41}$$

式中, $c_{l,\,N}>0$ 是常数.

将式 (39) ~ (41) 代入式 (38), 可得:

$$\begin{aligned}\dot{V}_{l,\,N} \le& -\sum_{j=1}^{N} c_{l,\,j}s_{l,\,j}^2 + \sum_{j=1}^{N}\Bigg[\eta_{l,\,j}\Theta_{l,\,j}^* + \eta_{l,\,j}\tilde{\Theta}_{l,\,j}\hat{\Theta}_{l,\,j} + \\ &\eta_{l,\,j}\tilde{\rho}_{l,\,j}\hat{\rho}_{l,\,j} - s_{l,\,1}\sum_{q=1}^{M} a_{l,\,q}\alpha_{q,\,2}^c\Bigg] + |s_{l,\,N}|\rho_{l,\,N}^* + \\ &s_{l,\,N}(u_l^a - \alpha_{l,\,N}^a) - \frac{s_{l,\,N}^2}{\sqrt{s_{l,\,N}^2+\eta_{l,\,N}^2}} + \\ &S_l^{\mathrm{T}}\left(\begin{bmatrix} h_{l,\,1}(s_{l,\,1}) \\ \vdots \\ h_{l,\,N}(s_{l,\,N})\end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & 1 & 0 \\ 0 & \cdots & 0 & 1\end{bmatrix} U_l^c\right)\end{aligned} \tag{42}$$

式中,$U_l^c = [\alpha_{l,\,1}^c,\ \alpha_{l,\,2}^c,\ \cdots,\ \alpha_{l,\,N-1}^c,\ u_l^c]$, $S_l^{\mathrm{T}} = [s_{l,\,1},\ \cdots,\ s_{l,\,N}]$.

为了减少执行器的更新次数, 设计如下的事件触发机制

$$\begin{cases} u_l^a(t) = \omega_l(t_k), & \forall t\in[t_k,\ t_{k+1}) \\ t_{k+1} = \inf\{t\in\mathbf{R} \mid |e_l(t)| \ge \delta|u_l^a(t)| + m\} \end{cases} \tag{43}$$

其中, $e_l(t) = u_l^a(t) - \omega_l(t)$; $0<\delta<1$; $m>0$ 是已知的常数.

根据式 (43), 存在两个时变值 $\lambda_i(t)$, $i=1,\ 2$, 满足 $|\lambda_i(t)|\le 1$ 成立, 并有

$$\omega_l(t) = [1+\lambda_1(t)\delta]u_l^a(t) + \lambda_2(t)m \tag{44}$$

根据式 (43) 和 (44), 可得:

$$\begin{aligned}\omega_l(t) =& -(1+\delta)\left(s_{l,\,N}\alpha_{l,\,N}^a\tanh\left(\frac{s_{l,\,N}\alpha_{l,\,N}^a}{\sigma_n}\right) + \right.\\ &\left. m\tanh\left(\frac{\alpha_{l,\,N}^a m}{\sigma_n}\right)\right)\end{aligned} \tag{45}$$

$$\frac{s_{l,\,N}\omega_l(t)}{1+\lambda_1(t)\delta} \le \frac{s_{l,\,N}\omega_l(t)}{1+\delta} \tag{46}$$

$$\left|\frac{\lambda_2(t)m}{1+\lambda_1(t)\delta}\right| \le \frac{m}{1-\delta} \tag{47}$$

$$|s_{l,\,N}|\rho_{l,\,N}^* - \frac{s_{l,\,N}^2}{\sqrt{s_{l,\,N}^2+\eta_{l,\,N}^2}} \le \eta_{l,\,N}\rho_{l,\,N}^* \tag{48}$$

式中, $\lambda_1(t)$、$\lambda_2(t)\in[-1,\ 1]$; σ_n 为设计参数.

将式 (44) ~ (48) 代入式 (42), 可得:

$$\begin{aligned}\dot{V}_{l,\,N} \le& -\sum_{j=1}^{N} c_{l,\,j}s_{l,\,j}^2 + \sum_{j=1}^{N}\Bigg[\eta_{l,\,j}(\rho_{l,\,j}^* + \Theta_{l,\,j}^*) + \\ &\eta_{l,\,j}\tilde{\Theta}_{l,\,j}\hat{\Theta}_{l,\,j} + \eta_{l,\,j}\tilde{\rho}_{l,\,j}\hat{\rho}_{l,\,j} - s_{l,\,1}\sum_{q=1}^{M} a_{l,\,q}\alpha_{q,\,2}^c\Bigg] + \\ &S_l^{\mathrm{T}}\left(\begin{bmatrix} h_{l,\,1}(s_{l,\,1}) \\ \vdots \\ h_{l,\,1}(s_{l,\,N})\end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & 1 & 0 \\ 0 & \cdots & 0 & 1\end{bmatrix} U_l^c\right)\end{aligned} \tag{49}$$

从式 (49) 的最后一项可知, 所设计的分布式前馈控制器 ${U_l}^a = [\alpha_{l,\,1}^a,\ \alpha_{l,\,2}^a,\ \cdots,\ \alpha_{l,\,N-1}^a,\ u_l^a]^{\mathrm{T}}$ 不能保证被控多智能体渐近稳定. 因此, 在第 2.2 节中, 分布式误差反馈优化控制器 U_l^c 将实现下列误差非线性多智能体系统稳定.

$$\begin{aligned}\dot{S}_l =& \begin{bmatrix} h_{l,\,1}(s_{l,\,1}) \\ h_{l,\,2}(s_{l,\,2}) \\ \vdots \\ h_{l,\,N}(s_{l,\,N})\end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & 1 & 0 \\ 0 & \cdots & 0 & 1\end{bmatrix} U_l^c - \\ &S_l\begin{bmatrix} \sum\limits_{q=1}^{M} a_{l,\,q} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & 0 & 0 \\ 0 & \cdots & 0 & 0\end{bmatrix}\alpha_{q,\,2}^c\end{aligned} \tag{50}$$

2.2　事件触发分布式误差反馈优化控制

本节将给出误差非线性多智能体系统 (50) 的事件触发分布式误差反馈优化控制器设计.

将式 (50) 变为如下的状态空间表达式:

$$\dot{S}_l = \bar{H}_l(S_l) + G_lU_l^c - \bar{G}_qU_q^c \tag{51}$$

式中, $\bar{H}_l(S_l) = [h_{l,\,1}(s_{l,\,1}),\ \cdots,\ h_{l,\,N}(s_{l,\,N})]^{\mathrm{T}}$; $G_l = \mathrm{diag}\{1,\ \cdots,\ 1\}^{\mathrm{T}}$; $\bar{G}_q = \mathrm{diag}\{\sum_{q=1}^{M} a_{l,\,q},\ 0,\ \cdots,\ 0\}^{\mathrm{T}}$, $U_q^c = [\alpha_{q,\,2}^c,\ 0,\ \cdots,\ 0]^{\mathrm{T}}$

为了同时实现协同跟踪与优化控制. 因此, 基于微分图博弈理论, 针对智能体 l, 建立如下的分布式代价函数

$$J_l = \int_0^{\infty}\bar{r}_l(S_l(\tau),\ U_l^c(\tau),\ U_q^c(\tau))\mathrm{d}\tau \tag{52}$$

式中, $\bar{r}_l(S_l(\tau), U_l^c(\tau), U_q^c(\tau)) = S_l^{\mathrm{T}}Q_lS_l + (U_l^c)^{\mathrm{T}}R_{ll}$

$\times U_l^c + \sum_{q\in M_l}(U_q^c)^{\mathrm{T}}R_{lq}U_q^c$, $Q_l = Q_l^{\mathrm{T}} > 0$ 和 $R_{ll} = R_{ll}^{\mathrm{T}} > 0$, $R_{lq} = R_{lq}^{\mathrm{T}} > 0$ 为权重矩阵, M_l 为智能体的个数.

定义如下的值函数

$$V_l(S_l) = \int_t^{\infty} \bar{r}_l(S_l(\tau),\ U_l^c(\tau),\ U_q^c(\tau))\mathrm{d}\tau \tag{53}$$

根据微分图博弈理论, 得到如下的 l-玩家 $l = 1, \cdots, M$ 的目标

$$V_l^*(S_l) = \min_{U_l^c} \int_t^{\infty} \bar{r}_l(S_l(\tau),\ U_l^c(\tau),\ U_q^c(\tau))\mathrm{d}\tau \tag{54}$$

定义 1[24−25]**.** 策略集 $u_1^*,\ u_2^*,\ \cdots,\ u_M^*$ 是全局 Nash 均衡的一个解, 如果

$$J_l^* := J_l(u_l^*,\ u_{\Xi-l}^*) \le J_l(u_l,\ u_{\Xi-l}^*) \tag{55}$$

式中, $u_{\Xi-l}^*$ 为图 Ξ 中除智能体 l 外的所有智能体的控制策略集, 即 $u_{\Xi-l}^* = \{u_q^* | q\in M_l,\ l\neq q\}$.

对于 $t\in[t_k,\ t_{k+1})$, 定义如下的事件触发误差

$$\bar{e}_l(t) = \bar{S}_{l,\ k} - S_l(t) \tag{56}$$

式中, $\bar{S}_{l,\ k} = S_l(t_k)$.

为了减少控制器 U_l^c 的频繁更新次数, 定义如下的事件触发机制

$$t_{k+1} = \inf\left\{ k\in\mathbf{Z}^+ \middle| (t>t_k)\wedge \left\{\|e_l\|^2 + \|e_q\|^2 \ge \frac{4Q_l(S_l)}{3\kappa\mu\bar{\upsilon}_l^2 L_{\nabla\varphi}^2} - \frac{4(\bar{\varphi}_l + \frac{1}{\kappa\mu})}{3\bar{\upsilon}_l^2 L_{\nabla\varphi}^2}(\|\hat{\upsilon}_{q,\ p}\|^2 + \|\hat{\upsilon}_{l,\ p}\|^2)\right\}\right\} \tag{57}$$

式中, $L_{\nabla\varphi} > 0$; $\kappa > 0$ 和 $\mu > 0$ 是常数; $\bar{\upsilon}_l$ 和 $\bar{\varphi}_l$ 是未知常数; $\hat{\upsilon}_{q,\ p}$ 和 $\hat{\upsilon}_{l,\ p}$ 表示触发后的第 q 和 l 个智能体的权重估计值 $\hat{\upsilon}$.

根据所设计的事件触发机制, 误差多智能体系统可变为如下形式

$$\dot{S}_l = \bar{H}_l(S_l) + G_l U_{l,\ k}^c - \bar{G}_q U_{q,\ k}^c \tag{58}$$

式中, $U_{l,\ k}^c = U_l^c(t_k)$ 为事件触发优化反馈控制律.

定义如下的哈密顿函数

$$\begin{aligned} &H_l(S_l,\ V_l,\ U_{l,\ k}^c) = \\ &\quad Q_l(S_l) + (U_{l,\ k}^c)^{\mathrm{T}}R_{ll}U_{l,\ k}^c + \sum_{q\in M_l}(U_{q,\ k}^c)^{\mathrm{T}}R_{lq}U_{q,\ k}^c + \\ &\quad \nabla V_l^{\mathrm{T}}(S_l)(\bar{H}_l(S_l) + G_l U_{l,\ k}^c - \bar{G}_q U_{q,\ k}^c) \end{aligned} \tag{59}$$

根据式 (59), 对哈密顿函数进行求偏导 $\frac{\partial H_l(S_l,\ V_l,\ U_{l,\ k}^c)}{\partial U_{l,\ k}^c} = 0$, 设计如下的分布式优化控制器

$$U_{l,\ k}^{c*} = -\frac{1}{2}R_{ll}^{-1}G_l^{\mathrm{T}}\nabla V_l^*(S_l) \tag{60}$$

将式 (60) 代入 (59), 可得:

$$\begin{aligned} 0 = {}& Q_l(S_l) + (U_{l,\ k}^{c*})^{\mathrm{T}}R_{ll}U_{l,\ k}^{c*} + \sum_{q\in M_l}(U_{q,\ k}^{c*})^{\mathrm{T}}R_{lq}U_{q,\ k}^{c*} + \\ & \nabla V_l^{*\mathrm{T}}(S_l)(\bar{H}_l(S_l) + G_l U_{l,\ k}^{c*} - \bar{G}_q U_{q,\ k}^{c*}) \end{aligned} \tag{61}$$

式中, $V^*(0) = 0$.

通常, 求解分布式最优控制器问题即寻找哈密顿–雅可比–贝尔曼方程 (61) 的解析解, 但该方程求解困难. 为此, 本文提出一种在线神经网络强化学习算法来学习每个智能体的哈密顿–雅可比–贝尔曼方程 (61) 的解析解.

2.3 强化学习算法

采用神经网络对代价函数进行估计, 并假设

$$V_l^*(S_l) = (\upsilon_l^*)^{\mathrm{T}}\varphi_l(S_l) + \varrho_l(S_l) \tag{62}$$

式中, $(\upsilon_l^*)^{\mathrm{T}} = [\upsilon_{l1}^*,\ \upsilon_{l2}^*,\ \cdots,\ \upsilon_{lN}^*]\in\mathbf{R}^l$ 是理想参数矩阵; $\varphi_l(S_l)\in\mathbf{R}^l$ 为神经网络激活函数; 假设 $\|\varrho_l(S_l)\| \le \bar{\varrho}_l$, 并且 $\bar{\varrho}_l > 0$ 是常数.

对式 (62) 求梯度

$$\nabla V_l^*(S_l) = \frac{\partial V_l^*(S_l)}{\partial S_l} = \nabla\varphi_l^{\mathrm{T}}(S_l)\upsilon_l^* + \nabla\varrho_l(S_l) \tag{63}$$

式中, $\nabla\varphi_l^{\mathrm{T}}(S_l)$ 和 $\nabla\varrho_l(S_l)$ 分别是 $\varphi_l^{\mathrm{T}}(S_l)$ 和 $\varrho_l(S_l)$ 的梯度, 并假设 $\|\nabla\varphi_l^{\mathrm{T}}(S_l)\| \le \bar{\varphi}_l$.

将式 (63) 代入式 (60), 可得:

$$U_{l,\ k}^{c*} = -\frac{1}{2}R_{ll}^{-1}G_l^{\mathrm{T}}(\nabla\varphi_l^{\mathrm{T}}(\bar{S}_{l,\ k})\upsilon_l^* + \nabla\varrho_l(\bar{S}_{l,\ k})) \tag{64}$$

由于理想权重 υ_l^* 是未知的, 用下列不等式估计 $V_l^*(S_l)$

$$\hat{V}_l(S_l) = \hat{\upsilon}_l^{\mathrm{T}}\varphi_l(S_l) \tag{65}$$

式中, $\hat{\upsilon}_l$ 是 υ_l^* 的估计.

将式 (65) 代入式 (64), 可得:

$$\hat{U}_{l,\ k}^c = -\frac{1}{2}R_{ll}^{-1}G_l^{\mathrm{T}}\nabla\varphi_l^{\mathrm{T}}(\bar{S}_{l,\ k})\hat{\upsilon}_l \tag{66}$$

利用经验回放技术设计自适应律, 定义如下的哈密顿误差

$$\begin{aligned} e_l(t) = {}& \hat{\upsilon}_{l,\ k}^{\mathrm{T}}\omega + S_l^{\mathrm{T}}Q_l S_l + (\hat{U}_{l,\ k}^c)^{\mathrm{T}}R_{ll}\hat{U}_{l,\ k}^c + \\ & \sum_{q\in M_l}(\hat{U}_{q,\ k}^c)^{\mathrm{T}}R_{lq}\hat{U}_{q,\ k}^c \end{aligned} \tag{67}$$

式中, $\hat{\upsilon}_{l,\ k} = \hat{\upsilon}_l(t_k)$; $\omega = \nabla\varphi_l(S_l)(\bar{H}_l(S_l) + G_l\hat{U}_{l,\ k}^c - \bar{G}_q\hat{U}_{q,\ k}^c)$.

定义如下的全局误差函数

$$E_{l,\,HJB} = \frac{1}{2}\left(\frac{e_l^2(t)}{(\omega^{\mathrm{T}}\omega+1)^2} + \sum_{p=1}^{m_t}\frac{e_{l,\,p}^2(t_p)}{(\omega_p^{\mathrm{T}}(t_p)\omega_p(t_p)+1)^2}\right) \tag{68}$$

其中, $\omega_p = \nabla\varphi_l(S_{l,\,p})(\bar{H}_l(S_{l,\,p}) + G_l\hat{U}_{l,\,k}^c - \bar{G}_q\hat{U}_{q,\,k}^c)$, $S_{l,\,p}$ 表示触发后的误差; $e_{l,\,p}(t_p)$ 表示历史数据; m_t 为历史数据的个数.

为在放宽持续激励条件[16−17]的同时确保受控系统的渐近稳定性, 本文采用历史数据 $e_{l,\,p}(t_p)$ 与当前数据 $e_l(t)$ 共同设计权重自适应律. 因此, $\hat{\upsilon}_l$ 的调整算法设计如下:

$$\dot{\hat{\upsilon}}_l = -\alpha\left(\frac{\omega e_l(t)}{(\omega^{\mathrm{T}}\omega+1)^2} + \sum_{p=1}^{m_t}\frac{\omega_p e_l(t_p)}{(\omega_p^{\mathrm{T}}(t_p)\omega_p(t_p)+1)^2}\right) \tag{69}$$

式中, α 是学习率;

$$e_l(t_p) = \hat{\upsilon}_l^{\mathrm{T}}\omega_p + S_{l,\,p}^{\mathrm{T}}Q_lS_{l,\,p} + (\hat{U}_{l,\,k}^c)^{\mathrm{T}}R_{ll}\hat{U}_{l,\,k}^c + \sum_{q\in M_l}(\hat{U}_{q,\,k}^c)^{\mathrm{T}}R_{lq}\hat{U}_{q,\,k}^c$$

根据式 (69), $\dot{\tilde{\upsilon}}_l$ 为

$$\dot{\tilde{\upsilon}}_l = -N_1 + N_2 \tag{70}$$

式中,

$$N_1 = \alpha\left(\frac{\omega\omega^{\mathrm{T}}}{(\omega^{\mathrm{T}}\omega+1)^2} + \sum_{p=1}^{m_t}\frac{\omega_p\omega_p^{\mathrm{T}}}{(\omega_p^{\mathrm{T}}\omega_p+1)^2}\right)\tilde{\upsilon}_l$$

$$N_2 = \alpha\left(\frac{\omega(\varrho_{l,\,H}+Y_l)}{(\omega^{\mathrm{T}}\omega+1)^2} + \sum_{p=1}^{m_t}\frac{\omega_p(\varrho_{l,\,H}+Y_{l,\,p})}{(\omega_p^{\mathrm{T}}\omega_p+1)^2}\right)$$

$$\varrho_{l,\,H} = -(\bar{H}_l(S_l) + G_lU_l^* - \bar{G}_qU_q^*)^{\mathrm{T}}\nabla\varrho_l$$

$$Y_l = \Bigg\{\left[\bar{G}_q(\hat{U}_{q,\,p}^c - U_q^{c*}) + G_l(\hat{U}_{l,\,p}^c - U_l^{c*})\right]^{\mathrm{T}}\times [\nabla\varphi_l^{\mathrm{T}}(S_l)\upsilon_l^*] + (\hat{U}_{l,\,p}^c)^{\mathrm{T}}R_{ll}\hat{U}_{l,\,p}^c - (U_l^{c*})^{\mathrm{T}}R_{ll}U_l^{c*} + \sum_{l\in M_l}(\hat{U}_{l,p}^c)^{\mathrm{T}}R_{lq}\hat{U}_{l,p}^c - \sum_{l\in M_l}(U_l^{c*})^{\mathrm{T}}R_{lq}U_l^{c*}\Bigg\}$$

$$Y_{l,\,p} = \Bigg\{\left[\bar{G}_q(\hat{U}_{q,\,p}^c - U_q^{c*}) + G_l(\hat{U}_{l,\,p}^c - U_l^{c*})\right]^{\mathrm{T}}\times [\nabla\varphi_l^{\mathrm{T}}(S_{l,\,q})\upsilon_l^*] + (\hat{U}_{l,\,p}^c)^{\mathrm{T}}R_{ll}\hat{U}_{l,\,p}^c - (U_l^{c*})^{\mathrm{T}}R_{ll}U_l^{c*} + \sum_{l\in M_l}(\hat{U}_{l,\,p}^c)^{\mathrm{T}}R_{lq}\hat{U}_{l,\,p}^c - \sum_{l\in M_l}(U_l^{c*})^{\mathrm{T}}R_{lq}U_l^{c*}\Bigg\}$$

根据文献 [22], 假设未知参数 υ_l^* 与激活函数 $\varphi_l(S_l)$ 均有界, 即满足 $\|\varphi_l\| \le \bar{\varphi}_l$, $\|\upsilon_l^*\| \le \bar{\upsilon}_l$.

为确保闭环系统 (51) 渐近稳定, 本文引入补偿项 ϑ 以消除近似误差的影响. 据此, 针对系统 (1) 设计的控制律如下所示:

$$\hat{U}_{l,\,k} = -\frac{1}{2}R_{ll}^{-1}G_l^{\mathrm{T}}\nabla\varphi_l^{\mathrm{T}}(S_l)\hat{\upsilon}_{l,\,k} + \vartheta \tag{71}$$

式中, $\vartheta = -\bar{\kappa}_2\|S_l\|^2/(\bar{\kappa}_1 + S_l^{\mathrm{T}}S_l)$; $\bar{\kappa}_1 > 0$ 是一个常数; $\bar{\kappa}_2 \in \mathbf{R}^+$ 满足如下不等式

$$\bar{\kappa}_2\|S_l\|^2 \ge \frac{\bar{\kappa}_1 + S_l^{\mathrm{T}}S_l}{(\bar{\varphi}_l\bar{\upsilon}_l + \bar{\varrho}_l)\|G_l\|}\Bigg[\frac{1}{4m_0}\|R_{ll}^{-1}\|^2\|G_l^{\mathrm{T}}\|^2\times (\bar{\upsilon}_l\bar{\varphi}_l + \bar{\varrho}_l)^2 + \frac{1}{4}\|R_{ll}^{-1}\|^2\|G_l^{\mathrm{T}}\|^2\times \left(\frac{m_0^3}{2}\bar{\upsilon}_l^{\mathrm{T}}\bar{\upsilon}_l\bar{\varphi}_l^2 + \left(m_0 + \frac{1}{2}\right)\bar{\varrho}_l^2\right)\Bigg]$$

3　稳定性分析

定理 1. 针对非线性多智能体系统, 分布式前馈事件触发控制器 (45) 与分布式最优控制器 (71), 以及参数自适应律 (12)、(13)、(26)、(27)、(40)、(41) 和 (69), 具有如下性质:

1) 被控系统 (1) 渐近稳定, 且跟随者输出 y_l 能够跟踪领导者输出 y_d;

2) 实现了微分图博弈的 Nash 均衡, 即达到了全局最优控制目标.

证明. 1) 分两种情况讨论被控多智能体系统的渐近稳定性, 构造如下李雅普诺夫函数:

$$V_{HJB} = \sum_{l=1}^{M}[V_{l,\,N} + \Phi_l + V_l^*(S_l) + V_l^*(S_{l,\,k})] \tag{72}$$

式中, $\Phi_l = \frac{1}{2\alpha}\tilde{\upsilon}_l^{\mathrm{T}}\tilde{\upsilon}_l$, $\tilde{\upsilon}_l = \upsilon_l^* - \hat{\upsilon}_l$, $||\upsilon_l^*|| \le \upsilon_M$, $\upsilon_M > 0$.

情况 1. 控制器无触发时刻, 即 $t_p < t < t_{p+1}$.

根据式 (72), $\dot{\Phi}_l$ 为

$$\dot{\Phi}_l \le -\tilde{\upsilon}_l^{\mathrm{T}}\Lambda\tilde{\upsilon}_l + (m_t+1)\|\tilde{\upsilon}_l\|(\|Y_l\| + \|Y_{l,\,p}\| + \sigma_{Hm}) \tag{73}$$

式中, $\Lambda^{\mathrm{T}} = \Lambda > 0$ 为正定矩阵; σ_{Hm} 为 $\varrho_{l,\,H}$ 的界;

$$\|Y_l\| \le [\|\bar{G}_q\|\|\hat{U}_{q,\,k}^c - U_q^{c*}\| + \|G_l\|\|\hat{U}_{l,\,k}^c - U_l^{c*}\|]\times [B_{\nabla\varphi}\upsilon_M] + \|Y_0\|$$

$$Y_0 = (\hat{U}_{l,\,k}^c)^{\mathrm{T}}R_{ll}\hat{U}_{l,\,k}^c + (U_l^{c*})^{\mathrm{T}}R_{ll}U_l^{c*} + \sum_{q\in M_l}(\hat{U}_{q,\,k}^c)^{\mathrm{T}}R_{lq}\hat{U}_{q,\,k}^c + \sum_{q\in M_l}(U_q^{c*})^{\mathrm{T}}R_{lq}U_q^{c*}$$

式中, $B_{\nabla\varphi}$ 是未知常数.

利用杨氏不等式, 可得:

$$\dot{\Phi}_l \leq -(\lambda_{\min}(\Lambda)-\lambda_1)\|\tilde{v}_l\|^2+\lambda_2\|\tilde{v}_l\|+(m_l+1)\times \frac{\left[\|\hat{U}^c_{l,\,k}-U^{c*}_l\|^2+\|\hat{U}^c_{q,\,k}-U^{c*}_q\|^2\right]}{2} \tag{74}$$

式中, $\lambda_1=2(m_l+1)B^2_{v_p}v^2_M(\|\bar{G}_q\|^2+\|G_l\|^2)$, B_{v_p} 是未知常数; $\lambda_2=2(m_l+1)(\|Y_0\|+\sigma_{Hm}/2)$; $\lambda_{\min}(\Lambda)$ 为 Λ 的最小特征值.

从式 (72), 可得:

$$\dot{V}_{l,\,N}\leq -\sum_{j=1}^{N}c_{l,\,j}S^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\tilde{\Theta}^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\tilde{\rho}^2_{l,\,j}+\sum_{j=1}^{N}\eta_{l,\,j}\left(\Theta^*_{l,\,j}+\rho^*_{l,\,j}+\frac{1}{2}\hat{\Theta}^2_{l,\,j}+\frac{1}{2}\hat{\rho}^2_{l,\,j}\right)+S^{\mathrm{T}}_l(\bar{H}_l(S_l)+G_l\hat{U}^c_{l,\,k}-\bar{G}_q\hat{U}^c_{q,\,k}) \tag{75}$$

利用杨氏不等式, 可得:

$$S^{\mathrm{T}}_l(\bar{H}_l(S_l)+G_l\hat{U}^c_{l,\,k}-\bar{G}_q\hat{U}^c_{q,\,k})\leq \frac{3}{2}\|S_l\|^2+\frac{1}{2}\|\nabla V_l\|+\frac{\|G_l\|^2}{2}\|\hat{U}^c_{l,\,k}-U^{c*}_l\|^2+\frac{\|\bar{G}_q\|^2}{2}\|\hat{U}^c_{q,\,k}-U^{c*}_q\|^2 \tag{76}$$

将式 (76) 代入式 (75), 可得:

$$\dot{V}_{l,\,N}\leq -\sum_{j=1}^{N}(c_{l,\,j}-\frac{3}{2})s^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\tilde{\Theta}^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\tilde{\rho}^2_{l,\,j}+\frac{1}{2}\|\nabla V_l\|+\frac{\|\bar{G}_q\|^2}{2}\|\hat{U}^c_{q,\,k}-U^{c*}_q\|^2+\sum_{j=1}^{N}\eta_{l,\,j}\left(\Theta^*_{l,\,j}+\rho^*_{l,\,j}+\frac{1}{2}\Theta^2_{l,\,j}+\frac{1}{2}\rho^2_{l,\,j}\right)+\frac{\|G_l\|^2}{2}\|\hat{U}^c_{l,\,k}-U^{c*}_l\|^2 \tag{77}$$

根据式 (72), $\dot{V}^*_l(S_l)$ 为

$$\dot{V}^*_l(S_l)\leq -\frac{1}{2}\lambda_{\min}(P)\|\nabla V_l\|^2+\|G_l\|^2\|\hat{U}^c_{l,\,k}-U^{c*}_l\|^2\times \|\bar{G}_q\|^2\|\hat{U}^c_{q,\,k}-U^{c*}_q\|^2 \tag{78}$$

式中, P 是对称正定矩阵.

利用杨氏不等式, 可得:

$$\left\|\hat{U}^c_{l,\,k}-U^{c*}_l\right\|^2\leq \left\|\frac{R^{-1}_{ll}G_l}{2}\right\|\left[\frac{3}{2}v^2_M L^2_{\nabla\varphi}\|s_{l,\,k}\|^2+2(B_{\nabla\varphi}\|\hat{v}_{l,\,k}\|+B_{\nabla\varphi}v_M+B_{\nabla\sigma})^2\right] \tag{79}$$

$$\left\|\hat{U}^c_{q,\,k}-U^{c*}_q\right\|^2\leq \frac{\|R^{-1}_{qq}\bar{G}_q\|}{2}\left[\frac{3}{2}v^2_M L^2_{\nabla\varphi}\|s_{q,\,k}\|^2+2(B_{\nabla\varphi}\|\hat{v}_{q,\,k}\|+B_{\nabla\varphi}v_M+B_{\nabla\sigma})^2\right] \tag{80}$$

式中, $B_{\nabla\sigma}$ 是未知常数.

将式 (73)、(77) 和 (78) 代入 (72), 可得:

$$V_{HJB}\leq\sum_{j=1}^{M}\Bigg\{-\sum_{j=1}^{N}(c_{l,\,j}-\frac{3}{2})s^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\Theta^2_{l,\,j}-\sum_{j=1}^{N}\frac{\eta_{l,\,j}}{2}\tilde{\rho}^2_{l,\,j}-\frac{1}{2}\lambda_{\min}(P)\|\nabla V_l\|^2+\lambda_2\|\tilde{v}_l\|+\sum_{j=1}^{N}\eta_{l,\,j}\left(\Theta^*_{l,\,j}+\rho^*_{l,\,j}+\frac{1}{2}\hat{\Theta}^2_{l,\,j}+\frac{1}{2}\hat{\rho}^2_{l,\,j}\right)+\frac{1}{2}\|\nabla V_l\|-(\lambda_{\min}(\Lambda)-\lambda_1)\|\tilde{v}_l\|^2-(U^{c*}_l)^{\mathrm{T}}R_{ll}U^{c*}_l+F_l+B_v\|G_l-\bar{G}_q\|\eta_{l,\,j}-\left[Q_l(S_l)-\frac{3}{4}\kappa\mu v^2_M\|e_{S_{l,\,p}}\|^2-\xi^2(t)-(\kappa\mu B_{\nabla\phi}+1)\|v_{q,\,k}\|^2\right]\Bigg\} \tag{81}$$

式中, $\kappa=\max\{\|R^{-1}_{11}G_1\|,\cdots,\|R^{-1}_{MM}G_M\|\}$; $F_l=2(\kappa\mu B_{\nabla\phi}+1)(B_{\nabla\phi}v_M+B_{\nabla\phi})^2+G_mB^2_{\nabla\phi}$; B_v 表示未知常数; $\mu=G_m/2+(m_l+1)/2+G_m/\lambda_{\min}(P)+\frac{1}{2}$, $G_m=\max\{\|\bar{G}_q\|^2,\|G_l\|^2\}$; $\xi=\sqrt{\|e_l\|^2+\|e_q\|^2}$.

情况 2. 控制器处于触发时刻, 即 $t=t_{p+1}$. 考虑式 (22) 的微分形式

$$\Delta V=\sum_{i=1}^{M}\left[\Delta V_{l,\,N}+\Delta\Phi_l+\Delta V_l(S_l)+\Delta V^*_l(S_{l,\,p})\right] \tag{82}$$

利用情况 1 的结论, 可得:

$$\Delta V(t)=\Delta V(t_{p+1})-\Delta V(t^*_{p+1})\leq\phi(S_{l,\,p+1}-S_{l,\,p}) \tag{83}$$

其中, $\phi(\cdot)$ 是一个类 $\mathcal{K}$ 函数. 因此该被控多智能体系统是渐近稳定的.

结合情况 1 和情况 2, 考虑事件触发条件 (59) 与 (60). 当 $\dot{V}\leq 0$ 成立, 可知 V 是有界的, 即 $s_{l,\,k}$ 有界. 由于 $\dot{s}_{l,\,k}$ 是有界的, 即 $s_{l,\,k}\in L_2(0,+\infty)$, 故由 Barbalat 引理可知, 被控系统是半全局渐近稳定的.

2) 证明最优控制器 U^{c*}_l 可实现全局 Nash均衡

根据式 (52), 可得:

$$J_l(S_l,\ U_l^c,\ U_q^c) = \int_0^\infty \left(S_l^{\mathrm{T}} Q_l S_l + (U_l^c)^{\mathrm{T}} R_{ll} U_l^c + \sum_{q\in M_l} (U_q^c)^{\mathrm{T}} R_{lq} U_q^c \right) \mathrm{d}\tau \tag{84}$$

由于被控系统是渐近稳定的, 有 $\lim_{t\to\infty} S_l(t)=0$, 即 $V_{HJB}(S_l(\infty))=0$, J_l 可写为:

$$\begin{aligned} J_l(S_l,\ U_l^c,\ U_q^c) =& \int_0^\infty \Bigg(Q_{l,\ q}(S_l) + (U_l^c)^{\mathrm{T}} R_{ll} U_l^c + \\ & \sum_{q\in M_l} (U_q^c)^{\mathrm{T}} R_{lq} U_q^c \Bigg) \mathrm{d}t + V_{HJB}(S_l(0)) + \\ & \int_0^\infty \dot{V}_{HJB} \mathrm{d}t = \int_0^\infty (U_l^c - U_l^{c*})^{\mathrm{T}} R_{ll} (U_l^c - U_l^{c*}) \mathrm{d}t - \\ & \sum_{q\in M_l} a_{l,\ q} (U_q^c - U_q^{c*})^{\mathrm{T}} R_{lq} (U_q^c - U_q^{c*}) \mathrm{d}t + \\ & V_{HJB}(S_{l,\ q}(0)) + \int_0^\infty (U_l^{c*})^{\mathrm{T}} R_{ll} (U_l^c - U_l^{c*}) \mathrm{d}t - \\ & \sum_{q\in M_l} a_{l,\ q} (U_q^{c*})^{\mathrm{T}} R_{lq} (U_q^c - U_q^{c*}) \mathrm{d}t \end{aligned} \tag{85}$$

因此, 根据 $U_l^c = U_l^{c*}$ 和 $U_q^c = U_q^{c*}$, 可得:

$$J_l(S_l,\ U_l^{c*},\ U_q^{c*}) = V_{HJB}(S_l(0)) \tag{86}$$

将式 (86) 代入式 (85), 可得:

$$\begin{aligned} J_l(S_l,\ U_l^c,\ U_q^{c*}) =& \int_0^\infty (U_l^c - U_l^{c*})^{\mathrm{T}} R_{ll} (U_l^c - U_l^{c*}) \mathrm{d}t + \\ & J_l(S_l,\ U_l^{c*},\ U_q^{c*}) \end{aligned} \tag{87}$$

对于任意 U_l, 可得:

$$J_l(S_l,\ U_l^{c*},\ U_q^{c*}) \le J_l(S_l,\ U_l^c,\ U_q^{c*}) \tag{88}$$

根据定义 1 可知, 所设计的分布式优化控制器 U_l^c 能渐近达到 Nash 均衡的解. □

以下将证明该系统不会出现 Zeno 现象. 在其前馈控制过程中, 存在常数 $\mu>0$ 使得 $t_{k+1}-t_k>\mu$ 成立. 定义 $e_l(t)=u_l^a(t)-\omega_l(t)$, 可得:

$$\frac{\mathrm{d}}{\mathrm{d}t}|e_l| = \mathrm{sign}(e_l)\dot{e}_l \le |\dot{u}_l^a(t)| \tag{89}$$

由于被控系统的所有信号有界, 对式 (89) 的等号两边进行积分, 可得

$$\int_{t_k}^{t_{k+1}} |\dot{e}_l| \mathrm{d}t \le \int_{t_k}^{t_{k+1}} L\ \mathrm{d}t \tag{90}$$

根据式 (90), 当 $t\in[t_k,\ t_{k+1})$ 时, 事件间隔 $t_{k+1}-t_k \ge L > 0$, 因此, 可得该前馈控制过程中不存在 Zeno 现象的结论.

求 $\dot{e}_l = \dot{S}_{l,\ k} - \dot{S}_l = -\dot{S}_l$, 利用式 (79) 和 (80), 当 $t\in[t_k,\ t_{k+1})$ 时, 有:

$$\begin{aligned} &\|\bar{H}_l(S_l) + G_l(U_l^c - e_l) - \bar{G}_q(U_q^c - e_l)\| \le \\ &\qquad L_Z\|e_l\| + L_Z\|S_l\| \end{aligned} \tag{91}$$

式中, L_Z 是正常数.

根据式 (91), $\|e_l/S_l\|$ 满足如下不等式

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\frac{e_l}{S_l}\right\| \le \frac{\dot{S}_l}{S_l}\left(1+\frac{e_j}{S_l}\right) = L_Z\left(1+\frac{e_j}{S_l}\right)^2 \tag{92}$$

定义 Γ, 可得 $\dot{\chi} \le L_Z\chi^2$ ($\chi = 1+\Gamma$) 和 $\mathrm{d}t \ge \frac{\mathrm{d}\chi}{L_Z\chi^2}$, $t\in[t_k,\ t_{k+1})$. 因此, 执行间隔 τ_k 满足 $\tau_k \ge -1/L_Z\chi|_{t_k}^{t}$. 由于在 $t\in[t_k,\ t_{k+1})$ 上 $e_l(t)>0$, 故 $\Gamma>0$, 这意味着 $\tau_k>0$. 综上, 在误差反馈最优控制过程中不会发生 Zeno 现象.

设计参数选择规则如下:

1) 选择常数 $c_{l,\ k}>0$、$\sigma_{l,\ k}>0$ 和 $\bar{\sigma}_{l,\ k}>0$ 能够定义虚拟前馈控制器和参数自适应律;

2) 选择常数 $c_{l,\ N}>0$、$\sigma_{l,\ N}>0$、$m>0$ 和 $\bar{\sigma}_{l,\ N}>0$ 能够定义事件触发前馈控制器和参数自适应律;

3) 选择常数 $\bar{\kappa}_1>0$、$\bar{\kappa}_2>0$ 和 $\alpha>0$ 能够定义优化控制器和参数自适应律.

注 1. 以上设计参数的选取只是保证被控系统稳定性的充分条件. 然而, 从式 (81) 可以看出, 增大设计参数 $c_{l,\ k}>0$, 减小设计参数 $\sigma_{l,\ k}>0$ 和 $\bar{\sigma}_{l,\ k}>0$ 将导致控制增益变大. 因此, 在实际应用中, 需要考虑跟踪性能和控制能量之间的权衡.

注 2. 本文提出了一种事件触发神经网络自适应混合式优化控制方法, 设计了神经网络自适应前馈控制器. 所设计的神经网络自适应前馈控制器把被控非线性多智能体系统转化为一个半全局稳定系统与误差系统的和. 针对误差系统, 应用动态规划理论, 提出了误差反馈优化控制器, 从而形成了"前馈控制 + 误差反馈校正"的神经网络自适应事件触发优化控制设计方法. 所提出的混合式优化控制方法能够克服现有智能自适应优化控制方法[8−11] 的"维数灾"问题.

4 仿真验证

本节以一组无人水面艇 (unmanned surface vehicle, USV) 为例, 验证所提出的分布式最优控制方法的优越性与有效性.

考虑由 1 个虚拟领导者和 3 艘无人水面艇组成的多智能体系统. 该多智能体系统处于有向通信图结构 (见图 2) 下. 各无人水面艇的艏摇动力学描述如下[26]:

$$\dot{\eta}_l = R_l(\bar{\psi}_l)\nu_l \tag{93}$$

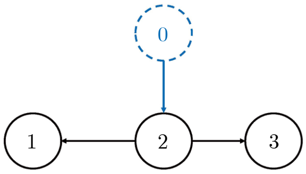

图 2 网络拓扑图

Fig. 2 Network topology diagram

$$\bar{G}_l\dot{\nu}_l=-C_l(\nu_l)\nu_l+u_l-D_l(\nu_l)\nu_l+d_l(t) \quad (94)$$

其中, $\eta_l=[\bar{\chi}_l,\ \bar{\phi}_l,\ \bar{\psi}_l]^{\rm T}$ 表示自主水面航行器的位置与艏摇角向量; $\nu_l=[\bar{\omega}_l,\ \bar{\upsilon}_l,\ \bar{r}_l]^{\rm T}$ 表示其纵荡、横荡及艏摇角速度; $u_l=[u_{1l},\ u_{2l},\ u_{3l}]^{\rm T}\in{\bf R}^3$ 为控制输入向量; $d_l(t)$ 表示由风、流、浪等因素引起的随机扰动; $\bar{G}_l\in{\bf R}^{3\times3}$ 为惯性矩阵; $C_l(\nu_l)$ 为科氏向心力矩阵; $D_l(\nu_l)$ 为非线性水动力阻尼矩阵; $R_l(\bar{\psi}_l)\in{\bf R}^{3\times3}$ 为旋转矩阵.

令 $x_{l,\,1}=\eta_l$ 和 $x_{l,\,2}=\nu_l$, 式 (93) 和 (94) 可变为:

$$\begin{cases}\dot{x}_{l,\,1}=R_lx_{l,\,2}\\ \dot{x}_{l,\,2}=\bar{G}_l^{-1}u_l+f_{l,\,2}(x_{l,\,2})+\bar{G}_l^{-1}d_l(t)\end{cases} \quad (95)$$

其中, $f_{l,\,2}(x_{l,\,2})=\bar{G}_l^{-1}[-C_l(\nu_l)\nu_l-D_l(\nu_l)\nu_l]$.

根据文献 [26], 通过一阶高斯−马尔科夫过程 $\dot{d}_l+\bar{\gamma}_ld_l=\bar{\delta}_l$ 构造外部扰动模型, 其中 $\bar{\gamma}_l$ 是参数, $\bar{\delta}_l$ 定义为白噪声.

定义如下的领导者动态

$$\begin{cases}\dot{\phi}_0=\begin{bmatrix}0&0.1&0\\-0.1&0&0\\0.01&0&0\end{bmatrix}\phi_0\\ y_d=\phi_0\end{cases} \quad (96)$$

式中, $\phi_0=[\phi_{1,\,0},\ \phi_{2,\,0},\ \phi_{3,\,0}]^{\rm T}$.

定义激活函数为 $\varphi_{l,\,k2}(x_{l,\,k2})$ 和 $\varphi_{l,\,k}(S_{l,\,k})$, $\varphi_{l,\,k2}(x_{l,\,k2})=[x_{l,\,k1}^2,\ x_{l,\,k1}x_{l,\,k2},\ x_{l,\,k2}^2]^{\rm T}$, $\varphi_{l,\,k}(S_{l,\,k})=[s_{l,\,k1}^2s_{l,\,k2},\ s_{l,\,k2}^2,\ s_{l,\,k1}^2,\ s_{l,\,k1}s_{l,\,k2}^2,\ s_{l,\,k1}^2s_{l,\,k2}^2]^{\rm T}$. 其中 $k=1,\ 2,\ 3$ 和 $l=1,\ 2,\ 3$.

根据文献 [25], 构造神经网络 $\upsilon_{l,\,k2}^{*\rm T}\varphi_{l,\,k}(x_{l,\,kd},\ x_{l,\,k2})$ 来逼近未知非线性函数 $F_{l,\,k2}(x_{l,\,k2})$, 构造神经网络 $\upsilon_{l,\,k}^{*\rm T}\varphi_{l,\,k}(s_{l,\,k1},\ s_{l,\,k2})$ 学习分布式代价函数 $J_l^*(s_{l,\,k1},\ s_{l,\,k2})$ 和分布式优化控制器 U_l^*.

选择 $Q_l=0.4$, $R_{ll}=0.5$ 和 $R_{lq}=0.5$, 分布式代价函数 (52) 为

$$J_l(S_l,\ U_l^c,\ U_q^c)=\int_0^\infty\left(0.4S_l^{\rm T}S_l+0.5U_l^{c\rm T}U_l^c+0.5\sum_{q\in M_l}U_q^{c\rm T}U_q^c\right){\rm d}\tau$$

在仿真中, 虚拟控制器 (11)、(25), 分布式前馈事件触发控制器 (45) 与分布式最优控制器 (71), 以及参数自适应律 (12)、(13)、(26)、(27)、(40)、(41) 和 (69) 的设计参数为 $c_{l,\,1}=25$, $\bar{\kappa}_1=2$, $c_{l,\,2}=30$, $\sigma_l=10$, $\sigma_{l,\,1}=0.30$, $\eta_{l,\,2}=0.45$, $\bar{\sigma}_{l,\,1}=0.33$, $\sigma_{l,\,2}=0.44$, $\alpha=5$, $\bar{\sigma}_{l,\,2}=0.30$, $\bar{\kappa}_2=3.22$, 可积函数选取为 $\eta_{l,\,1}={\rm e}^{-0.1t}$ 与 $\eta_{l,\,2}={\rm e}^{-0.2t}$.

选择初始条件为 $\eta_l(0)=[0.1,\ 0.2,\ 0.4]^{\rm T}$, $\upsilon_l(0)=[0.4,\ 0.2,\ 0.3]^{\rm T}$, $\hat{\Theta}_{l,\,1}(0)=0.2$, $\hat{\Theta}_{l,\,2}(0)=0.1$, $\hat{\rho}_{l,\,1}(0)=0.3$, $\hat{\rho}_{l,\,2}(0)=0.2$, $\hat{\upsilon}_l(0)=[0,\ 0,\ 0]^{\rm T}$.

该仿真结果如图 3 ~ 图 5 所示. 图 3 展示了跟随者输出和领导者输出的轨迹. 图 4 呈现了估计误差 $s_{l,\,k}$ 的曲线. 图 5 显示了分布式优化控制器的曲线.

根据图 3 ~ 图 5 的结果分析可知, 本文提出的分布式自适应神经网络优化控制方法能够保证跟随者的输出跟踪领导者的输出.

5 结论

本文研究了具有未知动力学的非线性多智能体

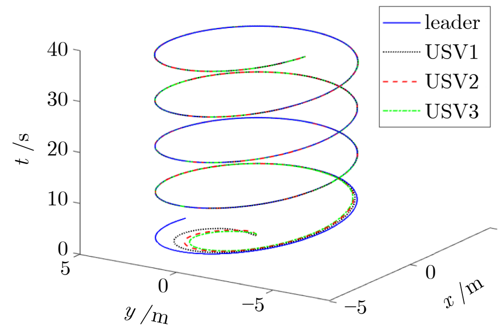

图 3 跟随者输出和领导者输出的轨迹

Fig. 3 Trajectories of follower output and leader output

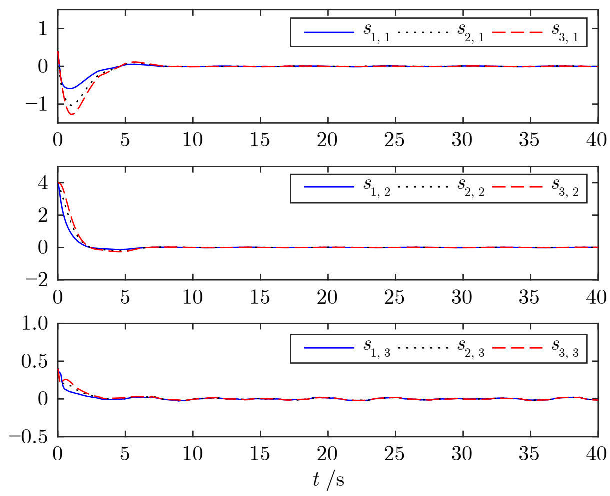

图 4 一致性跟踪误差的曲线

Fig. 4 The curves of consensus tracking errors

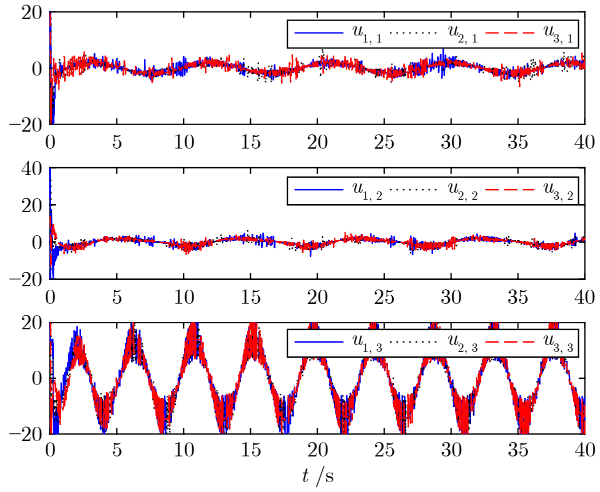

图 5 分布式优化控制器的曲线

Fig. 5 The curves of distributed optimal controllers

系统事件触发神经网络自适应分布式优化控制问题. 提出了一种事件触发分布式优化控制设计方法及相应的神经网络强化学习算法. 该算法能够学习分布式优化控制器与哈密顿–雅可比–贝尔曼方程的解析解, 其中评价神经网络权重的自适应更新律通过引入经验回放技术进行设计. 通过该神经网络学习算法获得的近似解可确保被控多智能体系统实现渐近稳定并达到 Nash 均衡. 仿真结果验证了所提分布式最优控制方法的有效性.

参考文献

1 He W L, Xu B, Han Q L, Qian F. Adaptive consensus control of linear multiagent systems with dynamic event-triggered strategies. *IEEE Transactions on Cybernetics*, 2020, **50**(7): 2996–3008

2 Li Z K, Liu X D, Ren W, Xie L H. Distributed tracking control for linear multiagent systems with a leader of bounded unknown input. *IEEE Transactions on Automatic Control*, 2013, **58**(2): 518–523

3 He W L, Zhang B, Han Q L, Qian F, Kurths J, Cao J D. Leader-following consensus of nonlinear multiagent systems with stochastic sampling. *IEEE Transactions on Cybernetics*, 2017, **47**(2): 327–338

4 Yu W W, Wang H, Cheng F, Yu X H, Wen G H. Second-order consensus in multiagent systems via distributed sliding mode control. *IEEE Transactions on Cybernetics*, 2017, **47**(8): 1872–1881

5 Wang F, Chen B, Lin C, Li X H. Distributed adaptive neural control for stochastic nonlinear multiagent systems. *IEEE Transactions on Cybernetics*, 2017, **47**(7): 1795–1803

6 Shen Q K, Jiang B, Shi P, Zhao J. Cooperative adaptive fuzzy tracking control for networked unknown nonlinear multiagent systems with time-varying actuator faults. *IEEE Transactions on Fuzzy Systems*, 2014, **22**(3): 494–504

7 Chen C L P, Ren C E, Du T. Fuzzy observed-based adaptive consensus tracking control for second-order multiagent systems with heterogeneous nonlinear dynamics. *IEEE Transactions on Fuzzy Systems*, 2016, **24**(4): 906–915

8 Zhang J L, Zhang H G, Feng T. Distributed optimal consensus control for nonlinear multiagent system with unknown dynamic. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(8): 3339–3348

9 Su W, Mu C X, Yang X, Bian J S, Sun C Y. Optimal tracking control for leader-following consensus of nonlinear multiagent systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2025, **55**(5): 3098–3107

10 Qin J H, Li M, Shi Y, Ma Q C, Zheng W X. Optimal synchronization control of multiagent systems with input saturation via off-policy reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, **30**(1): 85–96

11 Wen G X, Chen C L P, Feng J, Zhou N. Optimized multi-agent formation control based on an identifier-actor-critic reinforcement learning algorithm. *IEEE Transactions on Fuzzy Systems*, 2018, **26**(5): 2719–2731

12 Sun J L, Liu C S. Distributed fuzzy adaptive backstepping optimal control for nonlinear multimissile guidance systems with input saturation. *IEEE Transactions on Fuzzy Systems*, 2019, **27**(3): 447–461

13 Liu M S, Wan Y, Lopez V G, Lewis F L, Hewer G A, Estabridis K. Differential graphical game with distributed global Nash solution. *IEEE Transactions on Control of Network Systems*, 2021, **8**(3): 1371–1382

14 Qian Y Y, Liu M S, Wan Y, Lewis F L, Davoudi A. Distributed adaptive Nash equilibrium solution for differential graphical games. *IEEE Transactions on Cybernetics*, 2023, **53**(4): 2275–2287

15 Ren Y X, Wang Q S, Duan Z S. Optimal leader-following consensus control of multi-agent systems: A neural network based graphical game approach. *IEEE Transactions on Network Science and Engineering*, 2022, **9**(5): 3590–3601

16 Mazouchi M, Naghibi-Sistani M B, Sani S K H. A novel distributed optimal adaptive control algorithm for nonlinear multi-agent differential graphical games. *IEEE/CAA Journal of Automatica Sinica*, 2018, **5**(1): 331–341

17 Zhang Y W, Zhao B, Liu D R, Zhang S C. Distributed fault tolerant consensus control of nonlinear multiagent systems via adaptive dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(7): 9041–9053

18 Zhang H G, Zhang J L, Yang G H, Luo Y H. Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming. *IEEE Transactions on Fuzzy Systems*, 2015, **23**(1): 152–163

19 Wu W, Tong S C. Adaptive fuzzy distributed optimal FTC for nonlinear multiagent systems based multiplayer differential game. *IEEE Transactions on Fuzzy Systems*, 2025, **33**(2): 657–668

20 Long L J, Wang F L, Chen Z Y. Robust adaptive dynamic event-triggered control of switched nonlinear systems. *IEEE Transactions on Automatic Control*, 2023, **68**(8): 4873–4887

21 Cao L, Yao D Y, Li H Y, Meng W, Lu R Q. Fuzzy-based dynamic event triggering formation control for nonstrict-feedback nonlinear MASs. *Fuzzy Sets and Systems*, 2023, **452**: 1–22

22 Tong S C, Zhou H D, Li Y M. Neural network event-triggered formation fault-tolerant control for nonlinear multiagent systems with actuator faults. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(12): 7571–7582

23 Wei Q L, Jiao S S, Dong Q, Wang F Y. Event-triggered robust parallel optimal consensus control for multiagent systems. *IEEE/CAA Journal of Automatica Sinica*, 2025, **12**(1): 40−53

24 Zhang C J, Ji L H, Yang S D, Guo X, Li H Q. Distributed optimal consensus control for multiagent systems based on event-triggered and prioritized experience replay strategies. *Science China Information Sciences*, 2025, **68**(1): Article No. 112206

25 Lewis F L, Jagannathan S, Yesildirak A. *Neural Network Control of Robot Manipulators and Non-linear Systems*. London: CRC Press, 2020.

26 Peng Z H, Jiang Y, Liu L, Shi Y. Path-guided model-free flocking control of unmanned surface vehicles based on concurrent learning extended state observers. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(8): 4729−4739

吴 畏 辽宁工业大学理学院内聘副教授. 主要研究方向为非线性控制, 自适应控制, 非线性多智能体系统, 自适应分布式优化控制.
E-mail: wuw2017@163.com
(**WU Wei** Internal associate professor at the College of Science, Liaoning University of Technology. His research interests include nonlinear control, adaptive control, nonlinear multi-agent systems, and adaptive distributed optimal control.)

李克文 辽宁工业大学理学院副教授. 主要研究方向为预设时间控制, 优化控制, 非线性系统的模糊控制和自适应控制.
E-mail: likewen2018@163.com
(**LI Ke-Wen** Associate professor at the College of Science, Liaoning University of Technology. His research interests include prescribed-time control, optimal control, fuzzy control and adaptive control for nonlinear systems.)

佟绍成 辽宁工业大学理学院教授. 主要研究方向为模糊系统理论, 智能控制, 自适应控制和优化控制. 本文通信作者.
E-mail: jztongsc@163.com
(**TONG Shao-Cheng** Professor at the College of Science, Liaoning University of Technology. His research interests include fuzzy system theory, intelligent control, adaptive control, and optimal control. Corresponding author of this paper.)