

# 基于事件触发策略的多智能体系统的最优主-从一致性分析\*

刘 晨, 刘 磊

(河海大学 理学院, 南京 211100)

**摘要:** 研究了具有领导者的线性多智能体系统的主-从一致性问题,借助各智能体间的通讯拓扑所构成的无向图,提出一种基于事件触发的自适应动态规划方法,并使用神经网络的逼近性质设计出了近似最优控制,利用 Lyapunov 稳定性定理,分析了多智能体误差系统的稳定性,并找到一个该误差系统最终有界的充分条件,数值仿真结果进一步验证了理论分析的有效性。

**关键词:** 多智能体系统; 主-从一致性; 事件触发; 自适应动态规划

**中图分类号:** TP273; O232      **文献标志码:** A      **DOI:** 10.21656/1000-0887.400216

## 引 言

近年来,多智能体系统的协调追踪控制在生物系统集聚现象<sup>[1]</sup>、传感器网络<sup>[2]</sup>、无人飞行器编队控制<sup>[3]</sup>等领域取得了广泛应用。主-从一致性是协调追踪控制中的一类典型问题,它是指所有跟随者在适当的控制输入下趋向于领导者状态<sup>[4]</sup>。因此,针对此类问题的研究受到了众多团队的密切关注<sup>[5-6]</sup>。

现有的研究结果往往仅关注一致性问题,但是最优化作为一个重要特性没有受到充分的关注,目前仅有少量的研究结果。众所周知,优化问题的解可以通过对 Hamilton-Jacobi-Bellman (HJB) 方程的求解获得。然而,对于非线性系统 HJB 方程的求解,至今依旧是一个很大的难点。动态规划(dynamic programming, DP)是一种非常有效的求解优化控制问题的方法,但是对于非线性系统,随着维数的增长,计算量会呈指数倍增长,使得该算法实现越来越不可能。为了克服这一困难,Werbos 首次提出自适应动态规划(adaptive dynamic programming, ADP)求解算法,受到了广泛关注,并被众多学者进一步地发展<sup>[7-11]</sup>。Abu-Khalaf 和 Lewis 提出一种基于神经网络的最优算法,得出了一类具有控制约束的非线性系统的最优控制<sup>[12]</sup>。Liu 等将 ADP 算法推广到具有无限时域的非线性离散系统上,并证明了算法的收敛性<sup>[13]</sup>。近年来,Zhang 及其合作者在开发 ADP 算法求解最优控制问题上取得了众多成果<sup>[14-15]</sup>。文献[14]针对具有约束控制的非线性离散系统,提出了一种新颖的迭代 ADP 算法并证明了收敛性。文献[15]针对连续时

\* 收稿日期: 2019-07-16; 修订日期: 2019-10-10

**基金项目:** 国家自然科学基金(面上项目)(61773152);中国博士后科学基金(2016M601698; 2017T100318);江苏省博士后科学基金(面上项目)(1701078B)

**作者简介:** 刘晨(1993—),男,硕士生(E-mail: liuchen.com.cn@hhu.edu.cn);

刘磊(1983—),男,副教授,博士生导师(通讯作者。E-mail: liulei\_hust@163.com)。

间系统的最优跟踪问题,设计了一种新的 ADP 在线算法。

由于现实环境的复杂性,高维的多智能体系统动力学行为一般很难精确获得.在多智能体系统一致优化的研究中,越来越多的学者开始将强化学习中的行为-批评(actor-critic, AC)网络架构和 ADP 算法整合<sup>[16-18]</sup>.行为-批评网络架构是一种不依赖于模型的算法,利用神经网络来近似动态规划中的最优值函数和最优控制策略,通过环境或者被控对象在不同阶段产生的奖励惩罚来影响批评网络和行为网络,从而来获得最优值函数和最优控制以满足最优性原理.文献[16]针对一类离散时间多智能体系统,使用强化学习方法,提出了一种基于数据驱动的 ADP 解决了最优一致问题.文献[17]提出了一种 identifier-actor-critic 强化学习算法,实现了非线性多智能体系统的主-从编队优化控制.文献[18]利用一种基于观测器的 ADP 方法,解决了非线性多智能体系统的最优协同跟踪控制问题。

为减少控制任务的执行次数,降低通讯资源和计算资源的消耗,事件触发控制方法逐渐成为了一种有效的控制模式.相比于传统的连续时间采样控制,事件触发控制提供了一种非周期策略用来传输系统状态和控制输入,即仅在违反事件触发条件时才对系统状态进行采样.近年来,事件触发控制在线性系统<sup>[19-20]</sup>、非线性系统<sup>[21-22]</sup>、多智能体系统<sup>[23]</sup>等领域取得了广泛的应用.值得一提的是,Dong 和 Sun 等在文献[22]中首次将事件驱动策略和 AC 算法整合,实现了连续时间系统具有约束条件的最优控制.然而,基于事件触发的多智能体系统最优一致性问题尚未被充分研究,仅有极少量的文献提及此类问题的重要性,本文的主要目的就是填补这一空白.基于上述文献的启发,本研究针对通讯拓扑结构已知的线性多智能体系统,将事件触发策略和 AC 网络架构整合,实现了多智能体系统最优主从一致性控制,并且利用 Lyapunov 方法,对误差系统的有界性和稳定性进行了分析。

## 1 问题描述

### 1.1 代数图论

在多智能体系统中,各智能体之间的通讯拓扑可以使用无向图描述<sup>[24]</sup>.现考虑多智能体系统的无向图为  $\varphi = (\nu, \xi, \hat{A})$ .其中  $\nu = \{0, 1, 2, \dots, N\}$  为节点集,0 表示领导者,其余表示跟随者. $\xi \subseteq \nu \times \nu$  为系统边界集合. $\hat{A} = (a_{ij})_{N \times N}$  为跟随者的邻接矩阵且  $a_{ij} \geq 0$ .当且仅当节点  $i$  与  $j$  之间存在一条边,即  $\tilde{e}_{ij} \in \xi$  时  $a_{ij} > 0$ , 否则  $a_{ij} = 0$ .因此节点  $i$  的相邻节点集可表示为  $N_i = \{j \in \nu \mid \tilde{e}_{ij} \in \xi\}$ . $\mathbf{B} = \text{diag}\{b_i, i = 1, 2, \dots, N\}$  为领导者的邻接矩阵.定义度矩阵为  $\mathbf{D} = \text{diag}\{d_i, i = 1, 2, \dots, N\}$ , 其中  $d_i = \sum_{j \in N_i} a_{ij}$ . $\mathbf{L} = \mathbf{D} - \hat{A}$  为 Laplace 算子.记  $\mathbf{H} = (\mathbf{L} + \mathbf{B}) \otimes \mathbf{I}_n$ . $\Xi$  为  $R^n$  的一个紧集。

### 1.2 模型描述

考虑如下带有领导者与跟随者的多智能体系统:

$$\begin{cases} \dot{\mathbf{x}}_0(t) = \mathbf{A}\mathbf{x}_0(t), \\ \dot{\mathbf{x}}_i(t) = \mathbf{A}\mathbf{x}_i(t) + \mathbf{u}_i(t), \end{cases} \quad (1)$$

式中,  $\mathbf{A} \in R^{n \times n}$  为常数矩阵,  $\mathbf{x}_i(t) \in R^n$  与  $\mathbf{u}_i(t) \in R^n$  分别为智能体  $i$  的状态与控制输入,  $\mathbf{x}_0(t)$  为领导者状态。

**定义 1** 针对多智能体系统(1), 给定任意的初始状态  $\mathbf{x}_i(0), i = 1, 2, \dots, N$ , 若满足  $\lim_{t \rightarrow \infty} \|\mathbf{x}_i(t) - \mathbf{x}_0(t)\| = 0$ , 则称该系统达到主-从一致。

为了研究最优协调跟踪控制问题,定义智能体  $i$  的局部跟踪误差  $\mathbf{e}_i$  为

$$\mathbf{e}_i(t) = \sum_{j \in N_i} a_{ij}(\mathbf{x}_i(t) - \mathbf{x}_j(t)) + b_i(\mathbf{x}_i(t) - \mathbf{x}_0(t)). \quad (2)$$

引入 Kronecker 乘积, 则多智能体系统的全局追踪误差为

$$\begin{aligned} \mathbf{e}(t) &= (\mathbf{L} \otimes \mathbf{I}_n) \mathbf{x}(t) - (\mathbf{L} \otimes \mathbf{I}_n) \mathbf{x}_0(t) + (\mathbf{B} \otimes \mathbf{I}_n) (\mathbf{x}(t) - \mathbf{x}_0(t)) = \\ &= ((\mathbf{L} + \mathbf{B}) \otimes \mathbf{I}_n) (\mathbf{x}(t) - \mathbf{x}_0(t)) = \\ &= \mathbf{H}(\mathbf{x}(t) - \mathbf{x}_0(t)), \end{aligned} \quad (3)$$

$\mathbf{e}(t) = [\mathbf{e}_1^T(t), \mathbf{e}_2^T(t), \dots, \mathbf{e}_N^T(t)]^T \in R^{Nn}$ ,  $\mathbf{x}(t) = [\mathbf{x}_1^T(t), \mathbf{x}_2^T(t), \dots, \mathbf{x}_N^T(t)]^T \in R^{Nn}$  分别表示全局误差与状态向量,  $\mathbf{x}_0 = \mathbf{1} \cdot \mathbf{x}_0 \in R^{Nn}$ ,  $\mathbf{1} = \underline{\mathbf{1}} \otimes \mathbf{I}_n \in R^{Nn \times n}$ ,  $\underline{\mathbf{1}}$  表示元素全为 1 的  $N$  维列向量,  $\mathbf{I}_n$  为  $n$  阶单位阵.

**注 1** 由于  $\mathbf{H}$  为正定阵, 因此  $\mathbf{e}(t) \rightarrow \mathbf{0}$  等价于  $\mathbf{x}(t) \rightarrow \mathbf{x}_0(t)$ , 这意味着多智能体系统达到主-从一致, 即所有跟随者都服从于领导者.

**定义 2** (允许控制) 若任意的控制输入  $\mathbf{u}_i$  不仅使得系统 (2) 稳定, 而且能够使局部性能指标有界, 则称  $\mathbf{u}_i$  为允许控制.

定义智能体  $i$  的局部性能指标为

$$V_i(\mathbf{e}_i(0)) = \int_0^\infty [Q_i(\mathbf{e}_i) + W_i(\mathbf{u}_i)] dt, \quad (4)$$

式中,  $Q_i(\mathbf{e}_i) = \mathbf{e}_i^T \mathbf{Q}_i \mathbf{e}_i$ ,  $\mathbf{Q}_i \in R^{n \times n}$ ,  $W_i(\mathbf{u}_i) = \mathbf{u}_i^T R_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T R_{ij} \mathbf{u}_j$ ,  $\mathbf{u}_j$  为相邻节点  $j$  的控制输入, 同时  $\mathbf{Q}_i > \mathbf{0}$ ,  $R_{ii} > 0$ ,  $R_{ij} \geq 0$ . 当且仅当节点  $j$  为节点  $i$  的相邻接点时  $R_{ij} > 0$ , 否则  $R_{ij} = 0$ . 相应地, 局部值函数定义为

$$V_i(\mathbf{e}_i(t)) = \int_t^\infty [Q_i(\mathbf{e}_i) + W_i(\mathbf{u}_i)] dt. \quad (5)$$

针对于智能体  $i$ , 目标是找到一个允许控制  $\mathbf{u}_i$ , 使得多智能体趋于一致同时局部值函数最小:

$$V_i^*(\mathbf{e}_i(t)) = \min_{\mathbf{u}_i} \int_t^\infty [Q(\mathbf{e}_i) + W(\mathbf{u}_i)] dt. \quad (6)$$

根据文献 [10] 可知, 智能体  $i$  的 Lyapunov 方程 (LE) 为

$$\begin{aligned} LE_i(\mathbf{V}_{e_i}, \mathbf{u}_i) &= \mathbf{V}_{e_i}^T [\mathbf{H}_i(\dot{\mathbf{x}}(t) - \dot{\mathbf{x}}_0(t))] + Q_i(\mathbf{e}_i) + W_i(\mathbf{u}_i) = \\ &= \mathbf{V}_{e_i}^T [\mathbf{H}_i(\mathbf{A}\mathbf{H}^{-1}\mathbf{e}(t) + \mathbf{u}(t))] + \mathbf{e}_i^T \mathbf{Q}_i \mathbf{e}_i + \mathbf{u}_i^T R_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T R_{ij} \mathbf{u}_j = 0. \end{aligned} \quad (7)$$

假设局部最优值函数  $\mathbf{V}_{e_i}^* = \partial V / \partial \mathbf{e}_i$  满足上述 LE 方程, 则依据 Bellman 最优性原理, 得出局部最优控制为

$$\mathbf{u}_i^*(t) = -\frac{1}{2} R_{ii}^{-1} ((l_{ii} + b_{ii}) \otimes \mathbf{I}_n) \mathbf{V}_{e_i}^*. \quad (8)$$

### 1.3 事件触发控制

针对于智能体  $i$ , 事件触发序列  $\{t_k^i\}_{k=0}^\infty = 0 < t_0^i < t_1^i < \dots < t_k^i < \dots, \lim_{k \rightarrow \infty} t_k^i = \infty$ , 定义如下:

$$t_{k+1}^i = \inf \{t: t > t_k^i, \|\mathbf{r}_i(t)\| - P_i \geq 0\}, \quad \forall k \in \mathbf{N}, \quad (9)$$

式中,  $P_i$  为触发阈值, 多智能体误差系统的事件触发误差  $\mathbf{r}(t) = [\mathbf{r}_1^T(t), \mathbf{r}_2^T(t), \dots, \mathbf{r}_N^T(t)]^T$ ,  $\mathbf{r}_i(t) = \mathbf{e}_i(t) - \hat{\mathbf{e}}_k^i$ ,  $\hat{\mathbf{e}}_k^i = \mathbf{e}_i(t_k)$ .  $\{t_s\} = \{t_k^i: i = 1, 2, \dots, N, k \in \mathbf{N}_+\}$  表示所有智能体的事件触发时间集合.

## 2 基于事件触发的近似最优控制

### 2.1 批评网络

基于批评网络的智能体  $i$  的最优值函数可写为

$$V_i^*(\mathbf{e}_i) = \mathbf{W}_{ci}^T \boldsymbol{\phi}_{ci}(\mathbf{h}_i(\mathbf{e}_i)) + \varepsilon_{ci}, \quad (10)$$

式中,  $\mathbf{h}_i(\mathbf{e}_i) = [\mathbf{e}_i^T, \mathbf{u}_i^T]^T$ ,  $\mathbf{W}_{ci} \in R^n$  为目标权重矩阵,  $\boldsymbol{\phi}_{ci} \in R^n$  为激活函数,  $\varepsilon_{ci}$  为结构误差, 上式关于  $\mathbf{e}_i$  求偏导为

$$\mathbf{V}_{e_i}^* = (\nabla \boldsymbol{\phi}_{ci})^T(\mathbf{h}_i(\mathbf{e}_i)) \mathbf{W}_{ci} + \nabla \varepsilon_{ci}. \quad (11)$$

**假设 1** 在批评网络中, 目标权重矩阵  $\mathbf{W}_{ci}$  与激活函数  $\boldsymbol{\phi}_{ci}$  及结构误差  $\varepsilon_{ci}$  都是有界的, 即  $\|\mathbf{W}_{ci}\| \leq W_{cM}$ ,  $\|\boldsymbol{\phi}_{ci}\| \leq \phi_{cM}$ ,  $|\varepsilon_{ci}| \leq \varepsilon_{cM}$ ; 同时激活函数与结构误差的梯度  $\nabla \boldsymbol{\phi}_{ci}$ ,  $\nabla \varepsilon_{ci}$  也都是有界的, 即  $\|\nabla \boldsymbol{\phi}_{ci}\| \leq \nabla \phi_{cM}$ ,  $\|\nabla \varepsilon_{ci}\| \leq \nabla \varepsilon_{cM}$ ,  $\nabla \phi_{cM}$ ,  $\nabla \varepsilon_{cM}$  为大于零的常数.

利用批评网络 (10) 的逼近性质来估计值函数, 则 LE 方程为

$$H_i(\mathbf{e}_i, \mathbf{u}_i, \mathbf{W}_{ci}) = \mathbf{W}_{ci}^T \nabla \boldsymbol{\phi}_{ci}(\mathbf{h}_i(\mathbf{e}_i)) \dot{\mathbf{e}}_i + \mathbf{e}_i^T \mathbf{Q}_{ii} \mathbf{e}_i + \mathbf{u}_i^T \mathbf{R}_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T \mathbf{R}_{ij} \mathbf{u}_j = e_{Hi}. \quad (12)$$

由于目标权重矩阵  $\mathbf{W}_{ci}$  未知, 因此批评网络的实际输出可表示为

$$\hat{V}_i(\mathbf{e}_i) = \hat{\mathbf{W}}_{ci}^T \boldsymbol{\phi}_{ci}(\hat{\mathbf{h}}_i(\mathbf{e}_i)), \quad (13)$$

式中  $\hat{\mathbf{W}}_{ci}$  为未知目标权重矩阵  $\mathbf{W}_{ci}$  的估计值,  $\hat{V}_i(\mathbf{e}_i)$  为估计目标值函数, 因此估计的 LE 方程为

$$H_i(\mathbf{e}_i, \mathbf{u}_i, \hat{\mathbf{W}}_{ci}) = \hat{\mathbf{W}}_{ci}^T \nabla \boldsymbol{\phi}_{ci}(\mathbf{h}_i(\mathbf{e}_i)) \dot{\mathbf{e}}_i + \mathbf{e}_i^T \mathbf{Q}_{ii} \mathbf{e}_i + \mathbf{u}_i^T \mathbf{R}_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T \mathbf{R}_{ij} \mathbf{u}_j = e_{ci}. \quad (14)$$

批评网络的目标是选择一个恰当的  $\hat{\mathbf{W}}_{ci}$ , 使得误差函数最小:

$$E_{ci} = \frac{1}{2} e_{ci}^2. \quad (15)$$

估计权重  $\hat{\mathbf{W}}_{ci}$  更新规则如下:

$$\begin{cases} \hat{\mathbf{W}}_{ci} = \mathbf{0}, & t_k^i \leq t < t_{k+1}^i, \\ \hat{\mathbf{W}}_{ci}^+ = \hat{\mathbf{W}}_{ci} - l_{ci} \mathbf{k}_i \left( \mathbf{k}_{i1}^T \hat{\mathbf{W}}_{ci} + \mathbf{e}_i^T \mathbf{Q}_{ii} \mathbf{e}_i + \mathbf{u}_i^T \mathbf{R}_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T \mathbf{R}_{ij} \mathbf{u}_j \right), & t = t_k^i, \end{cases} \quad (16)$$

式中,  $\|\mathbf{k}_{i1}\| = \|\nabla \boldsymbol{\phi}_{ci}(\mathbf{h}_i(\mathbf{e}_i)) \dot{\mathbf{e}}_i\| \leq 1/2$ ,  $\|\mathbf{k}_i\| = \|\mathbf{k}_{i1}/(\mathbf{k}_{i1}^T \mathbf{k}_{i1} + 1)^2\| \leq 1/2$ ,  $l_{ci}$  为智能体  $i$  的权重学习率, 且  $0 < l_{ci} < 1$ .

### 2.2 行为网络

基于行为网络的智能体  $i$  的最优控制输入可写为

$$\mathbf{u}_i^*(t) = \mathbf{W}_{ai}^T \boldsymbol{\phi}_{ai}(\mathbf{e}_i) + \varepsilon_{ai}. \quad (17)$$

**假设 2** 在行为网络中, 目标权重矩阵  $\mathbf{W}_{ai} \in R^{n \times n}$  与激活函数  $\boldsymbol{\phi}_{ai} \in R^n$  及结构误差  $\varepsilon_{ai} \in R^n$  都是有界的, 即

$$\|\mathbf{W}_{ai}\| \leq \|\mathbf{W}_a\| \leq W_{aM}, \quad \|\boldsymbol{\phi}_{ai}\| \leq \|\boldsymbol{\phi}_a\| \leq \phi_{aM}, \quad \|\varepsilon_{ai}\| \leq \|\varepsilon_a\| \leq \varepsilon_{aM};$$

同时激活函数的梯度  $\nabla \boldsymbol{\phi}_{ai}$  是有界的, 即  $\|\nabla \boldsymbol{\phi}_{ai}\| \leq \nabla \phi_{aM}$ ,  $\nabla \phi_{aM}$  为大于零的常数. 对任意  $\mathbf{e}_i \in \Xi$ , 激活函数  $\boldsymbol{\phi}_{ai}$  Leibniz 连续, 即

$$\|\boldsymbol{\phi}_{ai}(\mathbf{e}_i(t)) - \boldsymbol{\phi}_{ai}(\hat{\mathbf{e}}_k^i)\| \leq L_{ai} \|\mathbf{r}_i(t)\|, \quad L_{ai} > 0.$$

同时

$$\mathbf{u}^*(t) = [\mathbf{u}_1^{*T}(t), \mathbf{u}_2^{*T}(t), \dots, \mathbf{u}_N^{*T}(t)]^T, \quad \mathbf{W}_a = [\mathbf{W}_{a1}^T, \mathbf{W}_{a2}^T, \dots, \mathbf{W}_{aN}^T]^T,$$

$$\boldsymbol{\phi}_a = [\boldsymbol{\phi}_{a1}, \boldsymbol{\phi}_{a2}, \dots, \boldsymbol{\phi}_{aN}]^T, \quad \varepsilon_a = [\varepsilon_{a1}^T, \varepsilon_{a2}^T, \dots, \varepsilon_{aN}^T]^T.$$

引入 Hadamard 乘积, 可得

$$\mathbf{u}^*(t) = \mathbf{W}_a \odot \boldsymbol{\phi}_a + \boldsymbol{\varepsilon}_a = (\mathbf{W}_{ai}^T \boldsymbol{\phi}_{ai} + \boldsymbol{\varepsilon}_{ai})_{N \times 1}.$$

由于目标权重  $\mathbf{W}_{ai}$  未知, 因此行为网络的事件触发近似最优控制为

$$\hat{\boldsymbol{\mu}}_i(t_k) = \hat{\mathbf{W}}_{ai}^T \boldsymbol{\phi}_{ai}(\hat{\mathbf{e}}_k^i), \quad (18)$$

式中,  $\hat{\boldsymbol{\mu}}_i(t_k)$  和  $\hat{\mathbf{W}}_{ai}$  分别为估计近似最优控制和估计权重. 定义行为网络的误差函数为

$$e_{ai} = \hat{V}_i(\hat{\mathbf{e}}_k^i). \quad (19)$$

类似于批评网络, 行为网络则是选择一个恰当的  $\hat{\mathbf{W}}_{ai}$ , 使得误差函数最小:

$$E_{ai} = \frac{1}{2} e_{ai}^2. \quad (20)$$

基于神经网络的误差反向传播原理对权重进行更新, 更新梯度如下:

$$\frac{\partial E_{ai}}{\partial \hat{\mathbf{W}}_{ai}} = \frac{\partial E_{ai}}{\partial e_{ai}} \frac{\partial e_{ai}}{\partial \hat{\mathbf{h}}_i(\hat{\mathbf{e}}_k^i)} \frac{\partial \hat{\mathbf{h}}_i(\hat{\mathbf{e}}_k^i)}{\partial \hat{\boldsymbol{\mu}}_i(t_k)} \frac{\partial \hat{\boldsymbol{\mu}}_i(t_k)}{\partial \hat{\mathbf{W}}_{ai}} =$$

$$W_{ci, n+1} \nabla \boldsymbol{\phi}_{ai}(\mathbf{e}_i) [ \hat{\mathbf{W}}_{ci}^T \boldsymbol{\phi}_{ci}(\hat{\mathbf{h}}_i(\mathbf{e}_i)) ], \quad (21)$$

$W_{ci, n+1}$  表示与行为网络控制输入相关的批评权重的绝对值. 因此估计权重  $\hat{\mathbf{W}}_{ai}$  的更新规则如下:

$$\begin{cases} \dot{\hat{\mathbf{W}}}_{ai} = \mathbf{0}, & t_k^i \leq t < t_{k+1}^i, \\ \hat{\mathbf{W}}_{ai}^+ = \hat{\mathbf{W}}_{ai} - l_{ai} \frac{\partial E_{ai}}{\partial \hat{\mathbf{W}}_{ai}}, & t = t_k^i, \end{cases} \quad (22)$$

式中,  $l_{ai}$  为智能体  $i$  的权重学习率, 且  $0 < l_{ai} < 1$ .

由于批评网络和行为网络的权重都只在触发瞬间更新, 因此本文希望设计一个合适的控制器, 使得误差系统稳定. 上述两个网络的权重估计误差定义为

$$\tilde{\mathbf{W}}_{ci} = \hat{\mathbf{W}}_{ci} - \mathbf{W}_{ci}, \quad \tilde{\mathbf{W}}_{ai} = \hat{\mathbf{W}}_{ai} - \mathbf{W}_{ai}. \quad (23)$$

**注 2** 本文基于多智能体系统的内部通讯拓扑结构, 利用行为-批评网络研究了各个智能体的一致性. 因此相比于文献[22], 本研究充分考虑了邻接节点的控制信息, 解决了内部结构更加复杂的系统的一致性.

**定理 1** 考虑多智能体系统(1), 如果假设 1 和假设 2 成立, 同时批评网络与行为网络权重按式(16)、(22)更新, 则在如下的事件触发条件下多智能体的误差系统(3)是最终有界的:

$$\| \mathbf{r}_i(t) \|^2 \leq \| \mathbf{r}(t) \|^2 \leq P_i^2 = \frac{\lambda_{\min}(\mathbf{Q}_{ii})(1 - \beta)}{s_i L_a^2 \| \hat{\mathbf{W}}_a^T \|^2} \| \mathbf{e}_i \|^2 + \sigma \eta^{-\delta t}, \quad (24)$$

式中,  $0 < \beta < 1, \sigma > 0, \eta > 0, \delta > 0$  为设计参数.

**证明** 为了证明系统的有界性, 需要考虑下述两类情形.

**情形 1** 事件未被触发 ( $t_k^i < t \leq t_{k+1}^i$ )

针对智能体  $i$ , 考虑如下 Lyapunov 函数:

$$L_i(t) = L_{i1}(t) + L_{i2}(t) + L_{i3}(t), \quad (25)$$

式中,  $L_{i1}(t) = 1/l_{ci} \cdot \text{tr}(\tilde{\mathbf{W}}_{ci}^T \tilde{\mathbf{W}}_{ci})$ ,  $L_{i2}(t) = 1/l_{ai} \cdot \text{tr}(\tilde{\mathbf{W}}_{ai}^T \tilde{\mathbf{W}}_{ai})$ ,  $L_{i3}(t) = V_i^*(\mathbf{e}_i)$ . 其中  $\dot{L}_{i1}(t) = \dot{L}_{i2}(t) = 0$ , 同时

$$\dot{L}_{i3}(t) = \dot{V}_i^*(\mathbf{e}_i) = \frac{\partial V_i^*}{\partial \mathbf{e}_i} \dot{\mathbf{e}}_i = \mathbf{V}_{e_i}^{*T} [ \mathbf{H}_i(\mathbf{A}\mathbf{H}^{-1}\mathbf{e}(t) + \hat{\mathbf{u}}(t_s)) ] =$$

$$\mathbf{V}_{e_i}^{*T} [ \mathbf{H}_i(\mathbf{A}\mathbf{H}^{-1}\mathbf{e}(t) + \mathbf{u}^*(t)) ] - \mathbf{V}_{e_i}^{*T} [ \mathbf{H}_i(\mathbf{u}^*(t) - \hat{\boldsymbol{\mu}}(t_s)) ]. \quad (26)$$

利用式(7)、(13)、(18), 式(26)可化为

$$\begin{aligned}
\dot{L}_{i3}(t) &= \frac{\partial V_i^*}{\partial \mathbf{e}_i} \dot{\mathbf{e}}_i = - \left( \mathbf{e}_i^T \mathbf{Q}_{ii} \mathbf{e}_i + \mathbf{u}_i^T \mathbf{R}_{ii} \mathbf{u}_i + \sum_{j \in N_i} \mathbf{u}_j^T \mathbf{R}_{ij} \mathbf{u}_j \right) - \\
&\quad \left( (\nabla \phi_{ci})^T (\mathbf{h}_i(\mathbf{e}_i)) \mathbf{W}_{ci} + \nabla \mathcal{E}_{ci} \right)^T \times \\
&\quad \left[ \mathbf{H}_i(\hat{\mathbf{W}}_a \odot (\phi_a(\mathbf{e}) - \phi_a(\mathbf{e}_k)) - \tilde{\mathbf{W}}_a \odot \phi_a(\mathbf{e}) + \boldsymbol{\varepsilon}_a) \right] \leq \\
&\quad - \mathbf{e}_i^T \mathbf{Q}_{ii} \mathbf{e}_i + \frac{1}{2} \left\| (\nabla \phi_{ci})^T (\mathbf{e}_i) \mathbf{W}_{ci} + \nabla \mathcal{E}_{ci} \right\|^2 + \\
&\quad \frac{1}{2} \left\| \mathbf{H}_i(\hat{\mathbf{W}}_a \odot (\phi_a(\mathbf{e}) - \phi_a(\mathbf{e}_k)) - \tilde{\mathbf{W}}_a \odot \phi_a(\mathbf{e}) + \boldsymbol{\varepsilon}_a) \right\|^2 \leq \\
&\quad - \lambda_{\min}(\mathbf{Q}_{ii}) \left\| \mathbf{e}_i \right\|^2 + \nabla \phi_{cM} \mathbf{W}_{cM}^2 + (\nabla \mathcal{E}_{cM})^2 + s_i L_{aM}^2 \left\| \hat{\mathbf{W}}_a \right\|^2 \left\| \mathbf{r} \right\|^2 + \\
&\quad s_i \phi_{aM}^2 \left\| \tilde{\mathbf{W}}_a^T \right\|^2 + s_i \boldsymbol{\varepsilon}_{aM}^2.
\end{aligned}$$

上式中

$$s_i = \left\| \mathbf{H}_i \right\|^2 = n \sum_{j=1}^n (l_{ij} + b_{ij})^2,$$

$$\boldsymbol{\varepsilon}_{\text{total}} = \nabla \phi_{cM} \mathbf{W}_{cM}^2 + (\nabla \mathcal{E}_{cM})^2 + s_i \phi_{aM}^2 \left\| \tilde{\mathbf{W}}_a^T \right\|^2 + s_i \boldsymbol{\varepsilon}_{aM}^2,$$

则

$$\begin{aligned}
\dot{L}_{i3}(t) &\leq - \lambda_{\min}(\mathbf{Q}_{ii}) \left\| \mathbf{e}_i \right\|^2 + s_i L_{aM}^2 \left\| \hat{\mathbf{W}}_a \right\|^2 \left\| \mathbf{r} \right\|^2 + \boldsymbol{\varepsilon}_{\text{total}} \leq \\
&\quad - \beta \cdot \lambda_{\min}(\mathbf{Q}_{ii}) \left\| \mathbf{e}_i \right\|^2 + \sigma s_i L_{aM}^2 \left\| \hat{\mathbf{W}}_a \right\|^2 \eta^{-\delta t} + \boldsymbol{\varepsilon}_{\text{total}}. \tag{27}
\end{aligned}$$

因此,当  $\left\| \mathbf{e}_i \right\| > ((\sigma s_i L_{aM}^2 \left\| \hat{\mathbf{W}}_a \right\|^2 + \boldsymbol{\varepsilon}_{\text{total}}) / \beta \cdot \lambda_{\min}(\mathbf{Q}_{ii}))^{1/2}$  时,  $\dot{L}_i(t) \leq 0$ , 即系统在事件间隔期间是最终有界的.

**情形 2** 事件被触发 ( $t = t_k$ )

对于相同的 Lyapunov 函数:

$$L_i(t) = L_{i1}(t) + L_{i2}(t) + L_{i3}(t). \tag{28}$$

第一项的一阶差分为

$$\Delta L_{i1}(t) = 1/l_{ci} \cdot (\text{tr}(\tilde{\mathbf{W}}_{ci}^{T+} \tilde{\mathbf{W}}_{ci}^+) - \text{tr}(\tilde{\mathbf{W}}_{ci}^T \tilde{\mathbf{W}}_{ci})), \tag{29}$$

同时  $\tilde{\mathbf{W}}_{ci}^+ = \hat{\mathbf{W}}_{ci} - l_{ci} \mathbf{k}_i (\mathbf{k}_{i1}^T \hat{\mathbf{W}}_{ci} + \boldsymbol{\varepsilon}_{Hi})$ , 于是

$$\begin{aligned}
\text{tr}(\tilde{\mathbf{W}}_{ci}^{T+} \tilde{\mathbf{W}}_{ci}^+) - \text{tr}(\tilde{\mathbf{W}}_{ci}^T \tilde{\mathbf{W}}_{ci}) &= \\
\text{tr}((\tilde{\mathbf{W}}_{ci} - l_{ci} \mathbf{k}_i (\mathbf{k}_{i1}^T \tilde{\mathbf{W}}_{ci} + \boldsymbol{\varepsilon}_{Hi}))^T (\tilde{\mathbf{W}}_{ci} - l_{ci} \mathbf{k}_i (\mathbf{k}_{i1}^T \tilde{\mathbf{W}}_{ci} + \boldsymbol{\varepsilon}_{Hi}))) - \text{tr}(\tilde{\mathbf{W}}_{ci}^T \tilde{\mathbf{W}}_{ci}) &= \\
- l_{ci} \mathbf{k}_i^T \mathbf{k}_{i1} \left\| \tilde{\mathbf{W}}_{ci} \right\|^2 - l_{ci} \text{tr}(\boldsymbol{\varepsilon}_{Hi} \tilde{\mathbf{W}}_{ci}) + \frac{1}{4} l_{ci}^2 \left\| \mathbf{k}_{i1}^T \tilde{\mathbf{W}}_{ci} + \boldsymbol{\varepsilon}_{Hi} \right\|^2 &\leq \\
- \left( \frac{1}{2} l_{ci} - \boldsymbol{\varepsilon}_{HM}^2 - \frac{1}{8} l_{ci}^2 \right) \left\| \tilde{\mathbf{W}}_{ci} \right\|^2 + \frac{1}{4} l_{ci}^2 (1 + 2\boldsymbol{\varepsilon}_{HM}^2). &\tag{30}
\end{aligned}$$

因此

$$\Delta L_{i1}(t) = - \left( \frac{1}{2} - \frac{\boldsymbol{\varepsilon}_{HM}^2}{l_{ci}} - \frac{1}{8} l_{ci} \right) \left\| \tilde{\mathbf{W}}_{ci} \right\|^2 + \frac{1}{4} l_{ci} (1 + 2\boldsymbol{\varepsilon}_{HM}^2). \tag{31}$$

行为网络权重更新为  $\tilde{\mathbf{W}}_{ai}^+ = \tilde{\mathbf{W}}_{ai} - l_{ai} \mathbf{W}_{ci,n+1} \nabla \phi_{ai}(\mathbf{e}_i) [\hat{\mathbf{W}}_{ci}^T \phi_{ci}(\hat{\mathbf{h}}_i(\mathbf{e}_i))]$ , 则第二项一阶差分为

$$\begin{aligned}
\Delta L_{i2}(t) &\leq - \mathbf{W}_{ci,n+1} \phi_{aM}^2 \left\| \tilde{\mathbf{W}}_{ai} \right\|^2 + \mathbf{W}_{ci,n+1} \phi_{cM}^2 \left\| \hat{\mathbf{W}}_{ci} \right\|^2 + \\
&\quad l_{ai} \mathbf{W}_{ci,n+1}^2 (\nabla \phi_{aM})^2 \phi_{cM}^2 \left\| \hat{\mathbf{W}}_{ci} \right\|^2. \tag{32}
\end{aligned}$$

当  $t = t_k$  时,  $\Delta L_{i3}(t) = V_i^*(\mathbf{e}_i^+) - V_i^*(\mathbf{e}_i) = 0$ . 综上, 可得

$$\Delta L_i(t) = \Delta L_{i1}(t) + \Delta L_{i2}(t) + \Delta L_{i3}(t) \leq$$

$$-a_i \|\tilde{\mathbf{W}}_{ci}\|^2 - b_i \|\tilde{\mathbf{W}}_{ai}\|^2 + D_i, \quad (33)$$

式中

$$D_i = 2W_{ci,n+1}\phi_{cM}^2(1 + l_{ai}W_{ci,n+1}(\nabla\phi_{aM})^2)W_{cM}^2 + \frac{1}{4}l_{ci}(1 + 2\varepsilon_{HM}^2),$$

$$a_i = \frac{1}{2} - \frac{\varepsilon_{HM}^2}{l_{ci}} - \frac{1}{8}l_{ci} - 2l_{ai}W_{ci,n+1}(\nabla\phi_{aM})^2\phi_{cM}^2, \quad b_i = W_{ci,n+1}(\nabla\phi_{aM})^2.$$

即当  $\|\tilde{\mathbf{W}}_{ci}\| (D_i/a_i)^{1/2} = m_{ci}$  或  $\|\tilde{\mathbf{W}}_{ai}\| > (D_i/b_i)^{1/2} = m_{ai}$  时,  $\Delta L_i \leq 0$ , 系统在触发瞬间是最终有界的.

接下来证明近似最优控制为允许控制,由上述分析可得

$$\begin{aligned} \|\tilde{\mathbf{u}}_i\| &= \|\hat{\boldsymbol{\mu}}_i - \mathbf{u}_i^*\| = \\ &\|\hat{\mathbf{W}}_{ai}^T(\boldsymbol{\phi}_{ai}(\mathbf{e}_i) - \boldsymbol{\phi}_{ai}(\hat{\mathbf{e}}_k)) - \tilde{\mathbf{W}}_a^T\boldsymbol{\phi}_{ai}(\mathbf{e}_i) + \boldsymbol{\varepsilon}_{ai}\| \leq \\ &\hat{W}_{aM}L_{ai}P_i + \tilde{W}_{aM}\phi_{aM} + \varepsilon_{aM}. \end{aligned} \quad (34)$$

上式中,已知触发条件  $P_i$  随着时间的推移趋于零,同时  $\phi_{aM}, \varepsilon_{aM}$  可以取到接近于零的数,因此事件触发近似最优控制  $\hat{\boldsymbol{\mu}}_i(t_k)$  最终会趋向于最优控制  $\mathbf{u}_i^*$ , 即  $\hat{\boldsymbol{\mu}}_i(t_k)$  为允许控制.

**注3** 在事件触发瞬间,系统状态  $\mathbf{x}_i(t)$  与近似最优控制  $\boldsymbol{\mu}_i(t_k)$  都进行了更新,然后,近似最优控制  $\boldsymbol{\mu}_i(t_k)$  保持不变直到下次事件触发瞬间再次进行更新.上述理论分析也证明了触发条件(24)能够对系统(1)进行有效地控制,同时相比于连续时间控制器,也节省了通讯资源与计算资源.

### 3 数值仿真

本节给出了一个数值例子来说明上述理论结果的可行性和有效性.考虑带有 1 个领导者和 4 个跟随者的多智能体系统,图 1 为该多智能体系统的通讯拓扑图.

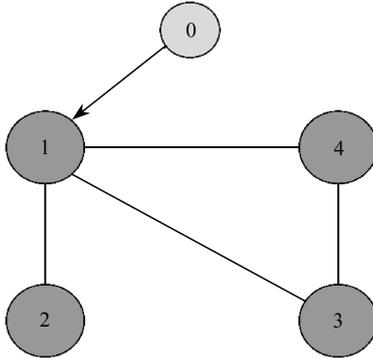


图 1 多智能体系统通讯拓扑图

Fig. 1 The communication topology of the multi-agent system

定义跟随者与领导者的邻接矩阵  $\hat{\mathbf{A}}$  与  $\mathbf{B}$  分别如下:

$$\hat{\mathbf{A}} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}.$$

不失一般性,令维数  $n = 1$ .进而假设该多智能体的系统矩阵  $A = -2$ ,领导者的初始状态为  $x_0(0) = 5$ ,第  $i$  个跟随者的初始状态为  $x_i(0) = 4 \times i, i = 1, 2, 3, 4$ .记控制时间  $T = 10$ ,每次更新

步长  $H = 0.001$ , 批评网络与行为网络中参数设置分别如下。

表 1 批评-行为网络参数取值

Table 1 Values of actor-critic network parameters

parameter	meaning	value
$l_{ci}$	learning rate of the critic network (CN)	0.01
$l_{ai}$	learning rate of the actor network (AN)	0.001
$N_c$	number of internal cycles of the CN	500
$N_a$	number of internal cycles of the AN	500
$T_{Ec}$	threshold for the CN	$e^{-10}$
$T_{Ea}$	threshold for the AN	$e^{-10}$
$N_{ch}$	number of hidden nodes in the CN	8
$N_{ah}$	number of hidden nodes in the AN	6

事件触发条件(24)中各参数取值分别为  $Q_{ii} = 1$ ,  $\beta = 0.7$ ,  $L_a = 2.5$ ,  $\sigma = 0.03$ ,  $\eta = 1.5$ ,  $\delta = 0.5$ 。基于此, 多智能体系统的数值仿真结果如图 2~4 所示。图 2 为多智能体系统中智能体 1 的事件触发阈值及状态误差的变化轨迹图。由图 2 可知, 在基于事件触发的自适应动态规划下, 节点 1 的误差最终趋于零。图 3 为多智能体系统中各智能体的状态变化轨迹图。由图 3 可知, 随着控制时间的推移, 4 个跟随者都与领导者达到最终一致。图 4 为基于神经网络设计出的近似最优控制的变化轨迹图。由图 4 可知, 在初期跟随者与领导者状态差别较大的情况下, 控制器作用较大, 在后期跟随者与领导者状态趋于一致后, 各个智能体的控制器均趋于零, 与图 2 结果互相吻合。

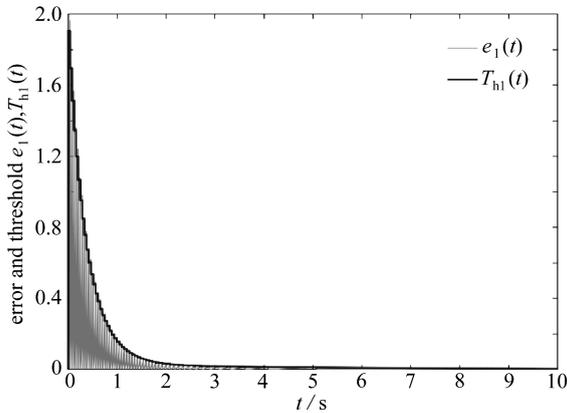


图 2 智能体 1 的误差及触发阈值

Fig. 2 Errors and trigger thresholds of agent 1

本研究针对多智能体系统设计了一种基于行为-批评网络的强化学习算法, 利用行为网络逼近最优控制, 利用批评网络逼近值函数, 并且这两个网络都只在事件触发瞬间更新。由图 2~4 可知, 该算法对模型取得了有效的控制, 并且相比于文献[19]中通过设计控制增益矩阵  $\mathbf{K}$  求解控制输入, 本文算法具有更高的容错率和资源节省率。

将事件触发条件(24)中的指数项  $\sigma\eta^{-\delta t}$  去除则变为文献[22]中的事件触发条件。利用文献[22]中的事件触发条件去控制上述案例, 运行出的数值仿真结果如图 5、6 所示。由图可知, 该多智能体系统需要 7~8 s 才能趋于主-从一致, 同时事件触发次数也有所增加。

事件触发阈值中的参数设置起到了至关重要的作用, 它控制着事件触发次数和收敛速度。

相比于文献[22]中的事件触发阈值,本研究的事件触发阈值函数添加了一项 $\sigma\eta^{-\delta t}$ ,因此通过设置参数 $\sigma, \eta, \delta$ 的大小能够更加灵活地控制触发次数及收敛速度.同时本研究中事件触发阈值 $P_i$ 始终大于0,这在一定程度上延长了事件触发间隔,也进一步验证了所提出的事件触发方法的有效性.

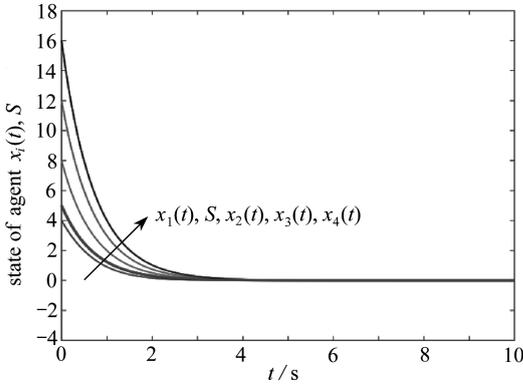


图3 各智能体的状态轨迹

Fig. 3 State trajectories of each agent

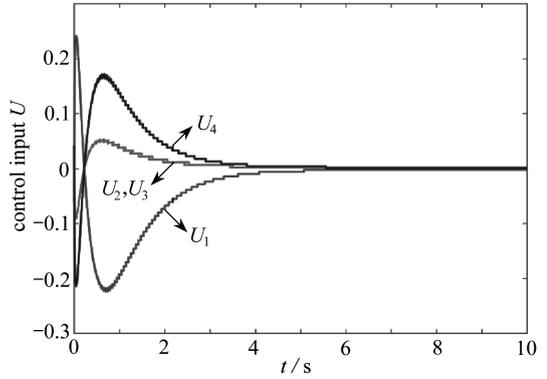


图4 各跟随者的优化控制

Fig. 4 Optimal control of each follower

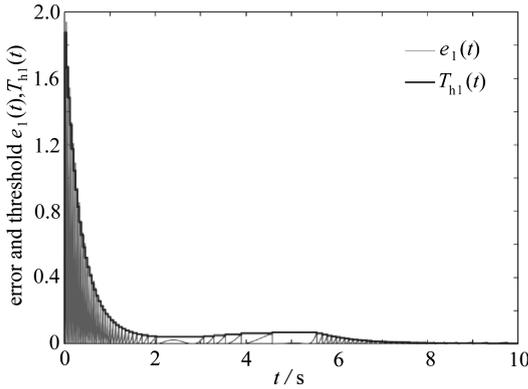


图5 智能体1的误差及触发阈值

Fig. 5 Errors and trigger thresholds of agent 1

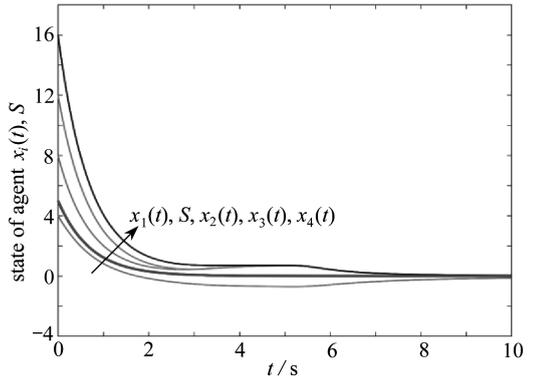


图6 各智能体的状态轨迹

Fig. 6 State trajectories of each agent

## 4 结 论

本研究基于事件触发的自适应动态规划方法,分析了多智能体的一致性.进而设计出了近似最优控制,证明了多智能体的误差系统最终有界,也验证了多智能体系统的主-从一致性.本文为后期研究带有控制约束的非线性多智能体系统的主-从一致性提供了一定的理论依据.

### 参考文献(References):

- [1] SU H, WANG X, LIN Z. Flocking of multi-agents with a virtual leader[J]. *IEEE Transactions on Automatic Control*, 2009, **54**(2): 293-307.
- [2] JI H, LEWIS F L, HOU Z, et al. Distributed information-weighted Kalman consensus filter for sensor networks[J]. *Automatica*, 2017, **77**: 18-30.
- [3] HU Q, ZHANG J. Relative position finite-time coordinated tracking control of spacecraft formation without velocity measurements[J]. *ISA Transactions*, 2015, **54**: 60-74.

- [4] LI H, LIAO X, HUANG T, et al. Event-triggering sampling based leader-following consensus in second-order multi-agent systems[J]. *IEEE Transactions on Automatic Control*, 2015, **60**(7): 1998-2003.
- [5] WEN G X, CHEN C L P, LIU Y J, et al. Neural-network-based adaptive leader-following consensus control for second-order non-linear multi-agent systems[J]. *IET Control Theory & Applications*, 2015, **9**(13): 1927-1934.
- [6] DEFOORT M, POLYAKOV A, DEMESURE G, et al. Leader-follower fixed-time consensus for multi-agent systems with unknown non-linear inherent dynamics[J]. *IET Control Theory & Applications*, 2015, **9**(14): 2165-2170.
- [7] BERTSEKAS D P, TSITSIKLIS J N. *Neuro-Dynamic Programming*[M]. Belmont, MA: Athena Scientific, 1996.
- [8] PROKHOROV D V, WUNSCH D C. Adaptive critic designs[J]. *IEEE Transactions on Neural Networks*, 1997, **8**(5): 997-1007.
- [9] MURRAY J J, COX C J, LENDARIS G G, et al. Adaptive dynamic programming[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2002, **32**(2): 140-153.
- [10] WANG F Y, ZHANG H, LIU D. Adaptive dynamic programming: an introduction[J]. *IEEE Computational Intelligence Magazine*, 2009, **4**(2): 39-47.
- [11] LI H, LIU D. Optimal control for discrete-time affine non-linear systems using general value iteration[J]. *IET Control Theory & Applications*, 2012, **6**(18): 2725-2736.
- [12] ABU-KHALAF M, LEWIS F L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach[J]. *Automatica*, 2005, **41**(5): 779-791.
- [13] LIU D, WEI Q. Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, **25**(3): 621-634.
- [14] ZHANG H, LUO Y, LIU D. Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints[J]. *IEEE Transactions on Neural Networks*, 2009, **20**(9): 1490-1503.
- [15] QIN C, ZHANG H, LUO Y. Online optimal tracking control of continuous-time linear systems with unknown dynamics by using adaptive dynamic programming[J]. *International Journal of Control*, 2014, **87**(5): 1000-1009.
- [16] ZHANG H, JIANG H, LUO Y, et al. Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method[J]. *IEEE Transactions on Industrial Electronics*, 2017, **64**(5): 4091-4100.
- [17] WEN G, CHEN C L P, FENG J, et al. Optimized multi-agent formation control based on an identifier-actor-critic reinforcement learning algorithm[J]. *IEEE Transactions on Fuzzy Systems*, 2018, **26**(5): 2719-2731.
- [18] ZHAO W, LI R, ZHANG H. Leader-follower optimal coordination tracking control for multi-agent systems with unknown internal states[J]. *Neurocomputing*, 2017, **249**: 171-181.
- [19] ZHU W, JIANG Z P. Event-based leader-following consensus of multi-agent systems with input time delay[J]. *IEEE Transactions on Automatic Control*, 2015, **60**(5): 1362-1367.
- [20] TAN X, CAO J, LI X, et al. Leader-following mean square consensus of stochastic multi-agent systems with input delay via event-triggered control[J]. *IET Control Theory & Applications*, 2017, **12**(2): 299-309.

- [21] TALLAPRAGADA P, CHOPRA N. On event triggered tracking for nonlinear systems[J]. *IEEE Transactions on Automatic Control*, 2013, **58**(9): 2343-2348.
- [22] DONG L, ZHONG X N, SUN C Y, et al. Event-triggered adaptive dynamic programming for continuous-time systems with control constraints[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, **28**(8): 1941-1952.
- [23] XU W, HO D W C, LI L, et al. Event-triggered schemes on leader-following consensus of general linear multiagent systems under different topologies[J]. *IEEE Transactions on Cybernetics*, 2015, **47**(1): 212-223.
- [24] 朱伟, 陈波. 具有领导者的非线性分数阶多智能体系统的一致性分析[J]. *应用数学和力学*, 2015, **36**(5): 555-562. (ZHU Wei, CHEN Bo. Leader-following consensus of fractional-order multi-agent systems with nonlinear models[J]. *Applied Mathematics and Mechanics*, 2015, **36**(5): 555-562. (in Chinese))

## Optimal Leader-Follower Consensus of Multi-Agent Systems Based on the Event-Triggered Strategy

LIU Chen, LIU Lei

(College of Science, Hohai University, Nanjing 211100, P.R.China)

**Abstract:** The leader-follower consensus of linear multi-agent systems was investigated. An event-triggered adaptive dynamic programming method was proposed based on the undirected graph formed by means of the communication topology among agents, and the approximate optimal control was designed with the approximate properties of neural networks. According to the Lyapunov stability theorem, the stability of multi-agent error systems was analyzed, and a sufficient condition for the ultimate boundedness of the error system was found. Finally, numerical simulation results further verify the effectiveness of the theoretical analysis.

**Key words:** multi-agent system; leader-follower consensus; event trigger; adaptive dynamic programming

**Foundation item:** The National Natural Science Foundation of China (General Program) (61773152); China Postdoctoral Science Foundation (2016M601698; 2017T100318)

---

引用本文/Cite this paper:

刘晨, 刘磊. 基于事件触发策略的多智能体系统的最优主-从一致性分析[J]. *应用数学和力学*, 2019, **40**(11): 1278-1288.

LIU Chen, LIU Lei. Optimal leader-follower consensus of multi-agent systems based on the event-triggered strategy[J]. *Applied Mathematics and Mechanics*, 2019, **40**(11): 1278-1288.