

基于滚动分布式鲁棒微分博弈的多智能体编队避碰控制

薛文艳^{1,2}, 黄捷^{2,3†}

(1. 广东海洋大学 机械与能源工程学院, 广东 阳江 529500;

2. 福州大学 电气工程与自动化学院, 福建 福州 350108; 3. 福州大学 5G+工业互联网研究院, 福建 福州 350108)

摘要: 针对未知障碍物环境下受有限通信能力以及干扰约束的多智能体编队控制存在鲁棒性差、安全性不佳的问题, 本文提出了一种新颖的滚动分布式鲁棒微分博弈方法. 首先, 与现有忽略外部干扰影响的流行分布式微分博弈方法比较, 将外部干扰看作虚拟的恶意参与者, 每个智能体基于局部交互信息优化最坏情况下的个体性能指标, 以实现对外部干扰的鲁棒性. 其次, 基于分布式终端状态估计器构造了一种近似开环鲁棒纳什均衡解, 以摆脱传统方法依赖全局状态信息求解的困境, 并从数学角度分析了近似纳什均衡的收敛性. 为保证编队的安全性, 在推导的近似纳什均衡解中引入避碰惩罚项, 并证明了系统的稳定性. 最后, 为了提升系统的安全性能, 基于滚动优化控制合成近似反馈鲁棒纳什策略, 通过对未知状态信息的预测, 减少了智能体在运动过程中死锁现象的发生. 仿真结果验证了所提方法的有效性.

关键词: 多智能体; 微分博弈; 近似纳什均衡; 反馈鲁棒纳什策略; 稳定性

引用格式: 薛文艳, 黄捷. 基于滚动分布式鲁棒微分博弈的多智能体编队避碰控制. 控制理论与应用, 2026, 43(4): 805 – 813

DOI: 10.7641/CTA.2024.40167

Multi-agent formation collision avoidance control with receding distributed robust differential game

XUE Wen-yan^{1,2}, HUANG Jie^{2,3†}

(1. College of Mechanical and Energy Engineering, Guangdong Ocean University, Yangjiang Guangdong 529500, China;

2. College of Electrical Engineering and Automation, Fuzhou University, Fuzhou Fujian 350108, China;

3. 5G+ Industrial Internet Institute, Fuzhou University, Fuzhou Fujian 350108, China)

Abstract: A novel receding distributed robust differential game method is proposed for the formation control problem of multi-agent systems under limited communication capabilities and disturbance constraints in unknown obstacle environments. Firstly, compared with existing popular distributed differential game methods that ignore external disturbance effects, external disturbance is regarded as virtual malicious participants, and each agent optimizes the worst-case individual performance indicators based on local interaction information to achieve robustness against external disturbance. Secondly, an approximate open-loop robust Nash equilibrium solution is constructed based on a distributed terminal state estimator to overcome the dilemma of relying on global state information in traditional methods, and theoretically ensure that the convergence of the approximate Nash equilibrium. Then, to ensure the safety of the formation, a collision avoidance penalty term is introduced into the derived approximate Nash equilibrium, and the stability of the system is proved. Finally, based on the receding horizon control, an approximate feedback robust Nash strategy is further synthesized to reduce the occurrence of deadlock by predicting unknown state information. The simulation results verify the effectiveness of the proposed method.

Key words: multi-agent; differential game; approximate Nash equilibrium; feedback robust Nash strategy; stability

Citation: XUE Wenyan, HUANG Jie. Multi-agent formation collision avoidance control with receding distributed robust differential game. *Control Theory & Applications*, 2026, 43(4): 805 – 813

收稿日期: 2024-03-22; 录用日期: 2024-11-11.

†通信作者. E-mail: jie.huang@fzu.edu.cn.

本文责任编辑: 贺威.

国家自然科学基金项目(92367109), 航空科学基金项目(20230001144001), 科研启动费项目(360302022401)资助.

Supported by the National Natural Science Foundation of China (92367109), the Aeronautical Science Foundation of China (20230001144001) and the Research Start-Up Fund (360302022401).

1 引言

多智能体系统(multi-agent systems, MASs)编队控制在工业、军事、娱乐等多个领域扮演着重要角色,例如水下探测^[1]、灯光表演^[2]以及协同围捕^[3]等.然而,在上述场景中, MASs需要解决由于智能体通信能力有限而无法获取全局信息的问题^[4].因此,研究分布式编队控制问题(distributed formation control problem, DFCEP)具有十分重要的意义,即需要设计某种分布式机制来解决受有限通信能力约束的MASs编队控制问题.

针对编队控制问题,已有学者基于微分博弈做出了令人瞩目的研究成果.文献[5]设计了一种非合作微分博弈模型,基于求解耦合黎卡提方程的方法得出各个参与者的纳什策略,实现了MASs的编队控制.然而,上述方法仅适用于集中式的场景,但是对于实际工作环境而言,智能体获取全局状态信息较困难.因此,为了解决多智能体之间的有限通信能力问题,文献[6]结合了分布式控制、最优控制以及博弈理论,提出一种多参与人分布式微分博弈方法,并在理论上分析了局部纳什均衡(local Nash equilibrium, L-NE)存在的条件.文献[7]考虑了有限时域和无限时域的微分博弈,根据智能体间的相互作用和几何属性,保证了L-NE轨迹在两种微分博弈模型下能够收敛到期望的队形,同时避免了邻居智能体之间的避碰.

在上述工作中,将微分博弈应用到编队问题中主要存在两方面问题:一方面,编队策略没有考虑障碍物约束,这对于在复杂环境中工作的多智能体通常是不现实的;另一方面,编队策略具有局部特性,即参与者仅能利用当前的局部信息获取纳什策略^[8].上述问题均会导致系统性能下降,使多智能体难以适应动态复杂的工作环境,甚至导致多智能体在动态复杂环境中发生碰撞.为此,文献[9]提出了一种滚动微分博弈方法,基于滚动优化控制,预测滚动窗口内未来一段时间的状态,在每次采样中,参与者仅选择开环NE(Nash equilibrium)生成的第1个控制输入作为决策变量,该过程在下次采样中重复进行,通过合成反馈NE策略来应对未知变化.然而,该NE策略没有考虑障碍物约束,使智能体难以适应存在复杂障碍物的编队环境.

另一个重要但经常被忽视的问题为不确定干扰对MASs的影响.不确定性干扰会影响参与者的性能指标,甚至可能导致个体之间发生碰撞^[10].为此,文献[11]引入了一种分布式鲁棒微分博弈(distributed robust differential game, DRDG)方法,将扰动视为虚拟的恶意参与者,试图最大化每个控制参与者的成本,并在强连通通信拓扑图的条件下,从理论上证明了局部鲁棒纳什均衡(local robust Nash equilibrium, LR-

NE)唯一存在性.然而,由于LR-NE的局部特性,参与者仅利用当前的状态信息更新博弈策略,使参与者在面对突发障碍物的出现时,容易陷入死锁状态,导致参与者无法继续前行,致使MASs安全性能下降.因此,为了提升系统的安全性能,文献[12]基于滚动优化控制与分布式状态观测器构建个体的近似纳什策略表达形式,形成了滚动分布式微分博弈(receding distributed differential game, RDDG)方法,该方法基于采样纳什技术合成在线近似反馈NE,在有向连通图的假设条件下保证了近似NE的存在性,为NE实现完全分布式奠定了理论基础.然而,该方法未考虑障碍物以及外界干扰约束,使参与者在具有障碍物和干扰存在的环境中适应性不佳,此外,近似NE的收敛性缺乏严格的理论证明.

为此,基于上述问题,本文针对未知障碍物环境下,受有限通信能力以及干扰约束的MASs编队控制问题,在文献[12]的基础上,提出了一种滚动分布式鲁棒微分博弈方法,以改善智能体的鲁棒性,增强智能体应对未知变化的适应能力,降低潜在的死锁风险.本文的主要贡献如下:

1) 将外部干扰因素看作恶意虚拟参与者,通过优化最坏情况下的性能指标,消解了干扰对个体运行的影响.进一步推导出仅依赖邻居终端状态信息的近似开环LR-NE,相比于文献[12]的RDDG方法,本文所提出的控制策略不仅提升了鲁棒性,还在理论上保证了近似LR-NE具有收敛性;

2) 相比于现有的仅考虑单控制目标的编队控制方法,例如文献[9, 12],本文综合考虑了编队与避碰目标,在推导的近似开环LR-NE中引入额外的避碰惩罚项,保证编队的安全性,并在理论上分析了MASs的稳定性;

3) 引入滚动优化控制,使智能体基于未来状态信息在线博弈,相比于笔者的前期工作(文献[11]),本文所提出的近似LR-NE策略减少了智能体运动过程中潜在死锁现象的发生,改善了智能体应对未知环境的能力,提升了系统的安全性能.

2 问题阐述

2.1 代数图论

考虑有向通信拓扑图为 $G(\mathcal{V}, \varepsilon)$,其中, $\mathcal{V} = \{1, \dots, i, \dots, N\}$ 表示图 G 的智能体集合,元素 $i \in \mathcal{V}$ 表示图 G 的第 i 个智能体; $\varepsilon \subseteq \mathcal{V} \times \mathcal{V}$ 表示图 G 的边集合;元素 $e_{ij} = (i, j) \in \varepsilon$ 表示第 i 个智能体可以获得第 j 个智能体的信息,即称第 j 个智能体是第 i 个智能体的邻居^[8].进一步,定义第 i 个智能体的邻居集合为 $\mathcal{N}_i \triangleq \{j | (i, j) \in \varepsilon\}$.若图中每一条边都有方向,且至少存在一个全局可达点,则称其为有向连通图.在本文中给出以下假设:

假设 1 智能体之间的通信拓扑图 $G(\mathcal{V}, \varepsilon)$ 是固定且有向连通的。

2.2 智能体的动态方程

MASs由 N ($N \geq 2$) 个智能体构成, 其中第 i 个智能体的动态方程如下:

$$\dot{x}_i = B_{ii}u_i + E_{ii}\omega_i, \quad \forall i \in N_{1:N}, \quad (1)$$

式中: $x_i(t) \in \mathbb{R}^3$ 表示第 i 个智能体的位置(状态信息); $u_i(t) \in \mathbb{R}^3$ 表示第 i 个智能体的控制输入; $B_{ii}(t) \in \mathbb{R}^{3 \times 3}$, $E_{ii}(t) \in \mathbb{R}^{3 \times 3}$ 分别表示第 i 个智能体的控制增益矩阵以及干扰增益矩阵, $B_{ii}(t)$ 是列满秩的; $\omega_i(t) \in \mathbb{R}^3$ 表示某种外部不确定的有界扰动。

假设外界干扰 $\omega_i(t) \in \mathbb{R}^3$ 为平方可积的, 即 $\forall t_f \in \mathbb{R}^+$, 存在常数 $\hat{p} \in \mathbb{R}^+$, 满足下式:

$$\int_{t_0}^{t_f} \|\omega_i(t)\|^2 dt < \hat{p}. \quad (2)$$

接着, 将多智能体系统的状态定义为 $x(t) \triangleq [x_1^T(t) \cdots x_i^T(t) \cdots x_N^T(t)]^T \in \mathbb{R}^{3N}$, 则MASs的动态方程为

$$\dot{x} \triangleq \sum_{i=1}^N (B_i u_i + E_i \omega_i), \quad (3)$$

式中: $B_i = [B_{1i}^T \cdots B_{Ni}^T]^T \in \mathbb{R}^{3N \times 3}$ 为控制增益矩阵; $E_i = [E_{1i}^T \cdots E_{Ni}^T]^T \in \mathbb{R}^{3N \times 3}$ 为干扰增益矩阵。

由于MASs受有限通信能力的约束, 因此, 本文引入分布式体系结构, 每个智能体基于局部交互信息计算控制策略. 进一步, 考虑无领导者-跟随者结构实现编队控制, 每个智能体仅能获取与邻居的期望相对编队位置. 定义第 i 个智能体的状态误差如下:

$$\tilde{x}_i(t) \triangleq \bar{A}_i(x(t) - \bar{x}^*) = \sum_{j \in \mathcal{N}_i} (x_i(t) - x_j(t) - d_{ij}^d), \quad (4)$$

式中: $\bar{x}^* \in \mathbb{R}^{3N}$ 表示MASs的期望状态; $d_{ij}^d \in \mathbb{R}^3$ 表示第 i 个智能体与第 j ($j \in \mathcal{N}_i$) 个智能体之间的期望相对编队位置; \bar{A}_i 表示第 i 个智能体的邻居信息矩阵(详见文献[11]).

定义 $\tilde{x} \triangleq [\tilde{x}_1^T \cdots \tilde{x}_i^T \cdots \tilde{x}_N^T]^T \in \mathbb{R}^{3N}$ 为多智能体系统的状态误差. 当且仅当系统的状态误差满足以下条件时, 每个智能体会实现期望的编队状态,

$$\lim_{t \rightarrow t_f} \tilde{x} = 0, \quad (5)$$

式中 t_f 表示系统执行编队任务的最终时间。

2.3 编队问题描述

本文考虑受外部不确定干扰以及障碍约束的网络化MASs的编队控制问题. 假设工作环境中存在 M 个已知障碍物和 \bar{N} 个未知障碍物, 其中总障碍物集合为 \mathcal{O} , 已知障碍物集合为 \mathcal{O} , 将环境中的障碍物膨化为椭

球形, \mathcal{S}_{ik} , \mathcal{A}_{ik} 和 \mathcal{M}_{ik} 分别表示第 i 个智能体与第 k 个障碍物构成的碰撞区域、感应区域以及自由区域^[11]. 为了便于DFCP的形成, 给出以下假设:

假设 2 1) 任意两个智能体的初始位置满足如下条件:

$$\|x_i(t_0) - x_j(t_0)\| \geq r_i + r_j, \quad \forall i \in N_{1:N}, i \neq j, \quad (6)$$

式中: r_i 表示第 i 个智能体的安全半径, r_j 表示第 j 个智能体的安全半径;

2) 每个智能体的初始位置与期望目标点位置分别满足以下条件:

$$x_i(t_0) \notin \mathcal{S}_{ik}, x_i(t_f) \in \mathcal{M}_{ik}, \quad \forall i \in N_{1:N}. \quad (7)$$

基于上述假设, 受外部干扰约束的网络化MASs的DFCP具体描述如下:

本文考虑由 N ($N \geq 2$) 个智能体组成的MASs, 其中每个智能体建模为式(1)–(4), 且每个智能体受外界干扰及有向通信拓扑图 $G(\mathcal{V}, \varepsilon)$ 约束. 将DFCP转化为滚动分布式鲁棒微分博弈(receding DRDG, R-DR-DG), 其中, 将执行编队任务的每个智能体看作参与人. 在采样时刻 t_k ($t_0 \leq t_k \leq t_{n-1}$), 每个参与人通过自身的效益函数 $J_{ik}^{\text{SH}}(x(t_k))$ 与邻居博弈, 个体通过最小化该效益函数获得最优鲁棒控制策略 u_{ik}^* 以及最坏干扰策略 ω_{ik}^* , 进一步地, 各个参与人无碰的实现期望编队任务, 最终, 第 i 个参与人与邻居参与人通过博弈组成的最优鲁棒策略集合 $\{(u_{ik}^*, \omega_{ik}^*), (u_{-ik}^*, \omega_{-ik}^*)\}$ 会收敛到LR-NE, 且该集合满足以下不等式组:

$$\begin{cases} J_{ik}^{\text{SH}*} \leq J_{ik}^{\text{SH}}(x(t_k), u_{ik}^*, \omega_{ik}^*, u_{-ik}^*, \omega_{-ik}^*), \\ J_{ik}^{\text{SH}*} \geq J_{ik}^{\text{SH}}(x(t_k), u_{ik}^*, \omega_{ik}^*, u_{-ik}^*, \omega_{-ik}^*), \end{cases} \quad (8)$$

式中: $J_{ik}^{\text{SH}*}$ 为 $J_{ik}^{\text{SH}}(x(t_k), u_{ik}^*, \omega_{ik}^*, u_{-ik}^*, \omega_{-ik}^*)$ 的缩写, 表示第 i 个参与人的LR-NE成本; u_{-ik}^* 表示第 i 个参与人的邻居对应的控制策略集合; ω_{-ik}^* 表示第 i 个参与人的邻居对应的干扰策略集合; u_{-ik}^* 与 ω_{-ik}^* 分别表示第 i 个参与人的邻居对应的最优控制策略与最坏干扰策略。

2.4 滚动分布式鲁棒微分博弈

本文所提出的R-DRDG方法求解近似LR-NE的原理图如图1所示, 在 $t \in [t_k, t_f]$ 内, 第 i ($\forall i \in N_{1:N}$) 个参与人将 $x_{ik}(t_k)$ 作为初始状态, 基于 J_{ik}^{SH} 与邻居非合作博弈. 第 i ($\forall i \in N_{1:N}$) 个参与人根据接收到的邻居终端状态 \hat{x}_{jk} 估计自身终端状态 \hat{x}_{ik} . 进一步将 \hat{x}_{ik} 发送给邻居. 进一步, 基于庞特里亚金最小值原理以及分布式终端状态估计器估计的终端状态推导出分布式鲁棒最优控制策略 u_{ik}^* 以及最坏干扰策略 ω_{ik}^* . 接着, 在 u_{ik}^* 中引入额外的势函数, 第 i ($\forall i \in N_{1:N}$) 个参与人通过与感应范围内的障碍物进行博弈, 规划出无碰轨迹. 最终, 第 i ($\forall i \in N_{1:N}$) 个参与人与邻居形成的最优鲁

棒策略集合 $\{(u_{i\bar{k}}^*, \omega_{i\bar{k}}^*), (u_{-i\bar{k}}^*, \omega_{-i\bar{k}}^*)\}$ 收敛到近似的开环LR-NE.

2.5 R-DRDG模型建立

首先, 将协作完成编队任务的多智能体看作博弈参与者. 其次, 设计各个参与者的博弈成本. 为了提升智能体适应未知环境的能力, 受文献[12]引入采样纳什思想的启发, 基于滚动优化控制理论, 构建如下形式的R-DRDG成本:

$$J_{i\bar{k}}^{\text{SH}} \triangleq \bar{\psi}(x_{\bar{k}}(t_f)) + \frac{1}{2} \int_{t_{\bar{k}}}^{t_f} (\|u_{i\bar{k}}\|_{R_i^u}^2 - \|\omega_{i\bar{k}}\|_{R_i^\omega}^2) dt, \quad (9)$$

式中: $J_{i\bar{k}}^{\text{SH}}$ 表示为 $J_{i\bar{k}}^{\text{SH}}(x(t_{\bar{k}}))$, $u_{i\bar{k}}, \omega_{i\bar{k}}, u_{-i\bar{k}}, \omega_{-i\bar{k}}$ 的

缩写; $t_{\bar{k}} \in [t_0, t_f]$ 表示为第 i 个参与者博弈过程的采样时刻, $t_0 < t_1 < \dots < t_{n-1} < t_f = t_n$, $\bar{k} = 0, 1, \dots, (n-1)$, $n \geq 1$; $\bar{\psi}(x_{i\bar{k}}(t_f)) = \frac{1}{2} \|x_{\bar{k}}(t_f) - \bar{x}^*\|_{F_i}^2$ 表示第 i 个参与者在 $t_{\bar{k}}$ 时刻的终端成本; $F_i = \bar{A}_i \bar{F}_i \bar{A}_i^T$; $\bar{F}_i = \bar{F}_i^T > 0$, $R_i^u = (R_i^u)^T > 0$ 与 $R_i^\omega = (R_i^\omega)^T > 0$ 均为正定可调节的权重矩阵.

注 1 文中所提到的死锁现象是指当智能体在躲避障碍物时, 可能会计算出一条避开障碍物的路径. 然而, 如果仅利用各个智能体获取的当前局部信息计算出的无碰路径, 可能导致它们无法沿着计划的路径移动而发生阻塞, 就会产生死锁现象. 在本文中, 只考虑即将到来的死锁, 而并没有涉及到直接死锁的避免.

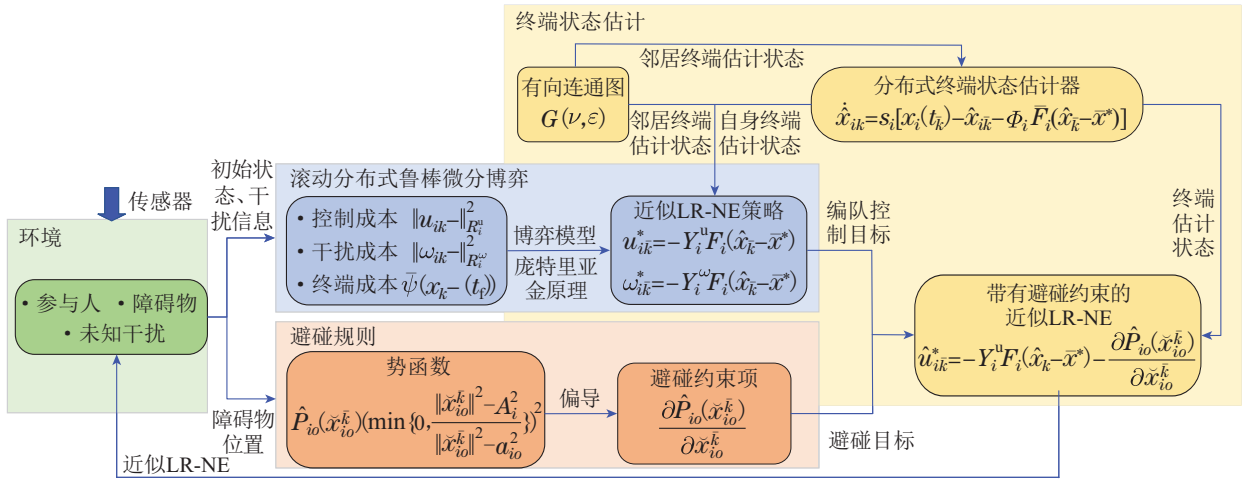


图1 近似LR-NE策略原理图

Fig. 1 The framework of approximate LR-NE strategy

2.6 LR-NE解的相关结论

以下定理旨在给出LR-NE解的表达形式, 并进一步分析该解的唯一存在性.

定理 1 对于由 N ($N \geq 2$) 个参与者组成的博弈系统, 每个参与者受式(1)–(4)以及网络拓扑结构 $G(\mathcal{V}, \varepsilon)$ 约束, 且对应的局部成本函数为式(9), 则每个参与者在采样时刻 $t_{\bar{k}} \in [t_0, t_f]$ 对应的LR-NE解是唯一存在的, 表达形式如下:

$$\begin{cases} u_{i\bar{k}}^* = -\Upsilon_i^u F_i \{(A)^{-1} [x(t_{\bar{k}}) + \sum_{i=1}^N \Phi_i F_i \bar{x}^*] - \bar{x}^*\}, \\ \omega_{i\bar{k}}^* = \Upsilon_i^\omega F_i \{(A)^{-1} [x(t_{\bar{k}}) + \sum_{i=1}^N \Phi_i F_i \bar{x}^*] - \bar{x}^*\}, \end{cases} \quad (10)$$

式中: 权重为 $\Upsilon_i^u = (R_i^u)^{-1} B_i^T$, $\Upsilon_i^\omega = (R_i^\omega)^{-1} E_i^T$, $\Phi_i = \int_{t_{\bar{k}}}^{t_f} [B_i (R_i^u)^{-1} B_i^T - E_i (R_i^\omega)^{-1} E_i^T] dt$, $A = I_{3N} + \sum_{i=1}^N \Phi_i F_i$.

该定理证明过程详见文献[12].

2.7 近似LR-NE解的分析

根据式(10)可知, 开环LR-NE解依赖MASs的当前状态 $x(t_{\bar{k}})$, 由于参与者之间受通信拓扑图约束, 因此, 无法获取所有参与者的状态信息. 为此, 基于庞特里亚金最小值原理, 通过一系列基本数学变换, 开环LR-NE解可以变形为

$$\begin{cases} u_{i\bar{k}}^* = -\Upsilon_i^u \bar{A}_i \bar{F}_i [\sum_{j \in \mathcal{N}_i} (x_{i\bar{k}}(t_f) - x_{j\bar{k}}(t_f) - d_{ij}^d)], \\ \omega_{i\bar{k}}^* = \Upsilon_i^\omega \bar{A}_i \bar{F}_i [\sum_{j \in \mathcal{N}_i} (x_{i\bar{k}}(t_f) - x_{j\bar{k}}(t_f) - d_{ij}^d)]. \end{cases} \quad (11)$$

基于式(11)可知, 参与者只需要获取自身与邻居的末端状态即可求解开环LR-NE.

在本小节, 首先为每个参与者设计分布式终端状态估计器; 其次给出近似开环LR-NE的收敛性分析; 最后, 引入额外的避碰约束, 以规划无碰的轨迹, 并分析系统的稳定性.

2.7.1 分布式终端状态估计器的设计

给出以下关于分布式终端状态估计器的定理.

定理2 考虑包含 $N (N \geq 2)$ 个参与人的模型, 其中每个参与人受式(1)–(4)以及网络拓扑结构 $G(\mathcal{V}, \varepsilon)$ 约束, 个体博弈成本满足式(9). 如果 $(-A)$ 矩阵是赫尔维茨矩阵, 且在采样时刻 $t_{\bar{k}} \in [t_0, t_f]$, 每个参与人根据以下分布式终端状态估计器更新对应的终端状态估计值 $\hat{x}_{i\bar{k}}$:

$$\dot{\hat{x}}_{i\bar{k}} = s_i \{x_i(t_{\bar{k}}) - \hat{x}_{i\bar{k}} - \Phi_i F_i(t_f) \bar{A}_i [\hat{x}_{\bar{k}} - \bar{x}^*]\}, \quad (12)$$

式中: s_i 表示分布式估计器增益; \bar{A}_i 表示第 $i (\forall i \in N_{1:N})$ 个参与人的邻居信息矩阵; $\hat{x}_{\bar{k}}(t) = [\hat{x}_{1\bar{k}}^T(t) \cdots \hat{x}_{N\bar{k}}^T(t) \cdots \hat{x}_{N\bar{k}}^T(t)]^T$ 表示在采样时刻 $t_{\bar{k}} \in [t_0, t_f]$, MASs的终端状态估计值. 且对应的近似LR-NE策略为

$$\begin{cases} \hat{u}_{i\bar{k}}^* = -\Upsilon_i^u F_i(t_f) [\hat{x}_{\bar{k}} - \bar{x}^*], \\ \hat{\omega}_{i\bar{k}}^* = \Upsilon_i^\omega F_i(t_f) [\hat{x}_{\bar{k}} - \bar{x}^*], \end{cases} \quad (13)$$

则每个参与人的近似LR-NE策略与真实LR-NE策略的偏差是指数收敛的.

证 基于式(11)与式(13), 可以推导近似LR-NE策略与真实LR-NE策略的偏差表达式形式如下:

$$\begin{cases} \Delta \hat{u}_{i\bar{k}} = -\Upsilon_i^u F_i(t_f) \Delta \hat{x}_{i\bar{k}}, \\ \Delta \hat{\omega}_{i\bar{k}} = \Upsilon_i^\omega F_i(t_f) \Delta \hat{x}_{i\bar{k}}, \end{cases} \quad (14)$$

式中 $\Delta \hat{x}_{i\bar{k}} = \hat{x}_{i\bar{k}}(t) - x_i(t_{\bar{k}})$.

进一步, 根据式(12)可以得出

$$x(t_{\bar{k}}) = x(t_f) + \sum_{i=1}^N \Phi_i F_i(t_f) [x_{\bar{k}}(t_f) - \bar{x}^*], \quad (15)$$

将式(15)代入式(12)中, 并经过基础数学变换, 可以推导出

$$\Delta \dot{\hat{x}}_{\bar{k}} = -SA \Delta \hat{x}_{\bar{k}}(t), \quad (16)$$

式中 $\Delta \hat{x}_{\bar{k}} = [\Delta \hat{x}_{1\bar{k}}^T \cdots \Delta \hat{x}_{i\bar{k}}^T \cdots \Delta \hat{x}_{N\bar{k}}^T]^T \in \mathbb{R}^{3N}$.

对式(16)求解, 并进一步对等式两边取欧几里得范数, 可以变形为

$$\|\Delta \hat{x}_{\bar{k}}(t)\| \leq \|\Delta \hat{x}_{\bar{k}}(t_{\bar{k}})\| e^{-\gamma_k t}, \quad (17)$$

式中 $\gamma_k = \min_m \operatorname{Re}[\lambda_m(SA)]$.

最后, 对式(14)等式两边取欧几里得范数, 并将式(17)代入, 可以得出 $\|\Delta \hat{u}_{i\bar{k}}(t)\| \leq \|\Upsilon_i^u F_i(t_f)\| \times \|\Delta \hat{x}_{i\bar{k}}(t_{\bar{k}})\| e^{-\gamma t}$, $\|\Delta \hat{\omega}_{i\bar{k}}(t)\| \leq \|\Upsilon_i^\omega F_i(t_f)\| \|\Delta \hat{x}_{i\bar{k}}(t_{\bar{k}})\| \times e^{-\gamma t}$, 其中, $\Delta u_{i\bar{k}}(t) = \hat{u}_{i\bar{k}}^* - u_{i\bar{k}}^*$, $\Delta \omega_{i\bar{k}}(t) = \hat{\omega}_{i\bar{k}}^* - \omega_{i\bar{k}}^*$. 因此, 每个参与人的近似LR-NE策略与真实LR-NE策略的偏差是指数收敛的. 证毕.

2.8 近似开环LR-NE的收敛性分析

给出以下关于近似LR-NE收敛性的定理.

定理3 考虑由 $N (N \geq 2)$ 个参与人组成的博弈系统, 其中每个参与人受式(1)–(4)以及网络拓扑结构 $G(\mathcal{V}, \varepsilon)$ 约束, 个体博弈成本满足式(9). 第 i 个参与

人与邻居组成的最优鲁棒控制策略集合 $\{(u_{i\bar{k}}^*, \omega_{i\bar{k}}^*), (u_{-i\bar{k}}^*, \omega_{-i\bar{k}}^*)\}$ 为开环近似LR-NE解, 且 $\epsilon_{\bar{k}}$ 满足以下形式:

$$\begin{cases} \epsilon_{\bar{k}} = \max_{i=1, \dots, N} \epsilon_{i\bar{k}}, \\ \epsilon_{i\bar{k}} = (\Pi_{\max}^{\bar{k}})^2 \|F_i(t_f)\| + \varrho_{\bar{k}} \|\Pi_{\bar{k}} \Psi_{\bar{k}} F_i(t_f)\| + \\ \quad \varrho_{\bar{k}}^2 \Xi_{i\bar{k}} + \varpi_{\bar{k}} \Theta_{i\bar{k}}, \\ \Pi_{\max}^{\bar{k}} = \max_{t_k \leq t \leq t_f} \|\Pi_{\bar{k}}(t)\|, \\ \Pi_{\bar{k}} = \sum_{i=1}^N \Phi_i F_i(t_f) \bar{x}^*, \\ \Psi_{\bar{k}} = \sum_{i=1}^N [B_i \Upsilon_i^u F_i(t_f) - E_i \Upsilon_i^\omega F_i(t_f)], \\ \Psi_{\max}^{\bar{k}} = \max_{t_k \leq t \leq t_f} \|\Psi_{\bar{k}}(t)\|, \\ \Xi_{i\bar{k}} = (\|F_i(t_f)\| + 4t_f \|\aleph_{i\bar{k}}^u\| + 4t_f \|\aleph_{i\bar{k}}^\omega\|) (\Psi_{\max}^{\bar{k}})^2, \\ \Theta_{i\bar{k}} = \|\aleph_{i\bar{k}}^u\| + \|\aleph_{i\bar{k}}^\omega\|, \\ \aleph_{i\bar{k}}^u = (\Upsilon_i^u)^T F_i^T(t_f) R_i^u \Upsilon_i^u F_i(t_f), \\ \aleph_{i\bar{k}}^\omega = (\Upsilon_i^\omega)^T F_i^T(t_f) R_i^\omega \Upsilon_i^\omega F_i(t_f). \end{cases} \quad (18)$$

证 令 $\tilde{J}_{i\bar{k}}^{\text{SH}}(x(t_{\bar{k}}), u_{i\bar{k}}, \omega_{i\bar{k}}, u_{-i\bar{k}}, \omega_{-i\bar{k}})$ 表示第 i 个参与人在采样时刻 $t_{\bar{k}} \in [t_0, t_f]$ 时, 使用近似的LR-NE解(13)对应的博弈成本, $\tilde{J}_{i\bar{k}}^{\text{SH}}(x(t_{\bar{k}}), u_{i\bar{k}}, \omega_{i\bar{k}}, u_{-i\bar{k}}, \omega_{-i\bar{k}})$ 表示第 i 个参与人在采样时刻 $t_{\bar{k}} \in [t_0, t_f]$, 使用真实的LR-NE解(11), 其他参与人使用近似的LR-NE解(13)对应的博弈成本. 进一步, 可以得出

$$\Delta J_{i\bar{k}}^{\text{SH}} = \frac{1}{2} \|\hat{x}_{\bar{k}}(t_f) - \bar{x}^*\|_{F_i}^2 - \frac{1}{2} \|x_{\bar{k}}(t_f) - \bar{x}^*\|_{F_i}^2 + \frac{1}{2} \int_0^{t_f} (\|\Delta \hat{u}_{i\bar{k}}\|_{R_i^u}^2 - \|\Delta \hat{\omega}_{i\bar{k}}\|_{R_i^\omega}^2) dt, \quad (19)$$

式中: $\tilde{J}_{i\bar{k}}^{\text{SH}}$ 表示 $\tilde{J}_{i\bar{k}}^{\text{SH}}(x(t_{\bar{k}}), u_{i\bar{k}}, \omega_{i\bar{k}}, u_{-i\bar{k}}, \omega_{-i\bar{k}})$ 的缩写; $\tilde{J}_{i\bar{k}}^{\text{SH}}$ 表示 $\tilde{J}_{i\bar{k}}^{\text{SH}}(x(t_{\bar{k}}), u_{i\bar{k}}, \omega_{i\bar{k}}, u_{-i\bar{k}}, \omega_{-i\bar{k}})$ 的缩写; $\Delta J_{i\bar{k}}^{\text{SH}} = \tilde{J}_{i\bar{k}}^{\text{SH}} - \tilde{J}_{i\bar{k}}^{\text{SH}}$.

如果每个参与人使用近似的LR-NE解(13), 则MASs的动力学方程满足

$$\dot{\hat{x}}_{\bar{k}} \triangleq -\Psi_{\bar{k}} \hat{x}_{\bar{k}}. \quad (20)$$

将式(20)代入式(19)中, 并对等式两边从 $t_{\bar{k}}$ 到 t_f 积分, 基于一系列数学变换可以推导出

$$\begin{aligned} & \frac{1}{2} \int_{t_{\bar{k}}}^{t_f} \|\Delta \hat{u}_{i\bar{k}}\|_{R_i^u}^2 dt = \\ & \frac{1}{2} \int_{t_{\bar{k}}}^{t_f} \|2 \int_{t_{\bar{k}}}^{t_f} \Psi_{\bar{k}} \nu_{\bar{k}}(t) dt - \nu_{\bar{k}}(t)\|_{R_i^u}^2 dt \leq \\ & 4t_f \left(\int_{t_{\bar{k}}}^{t_f} \Psi_{\bar{k}} \nu_{\bar{k}}(t) dt \right)^T \aleph_{i\bar{k}}^u \left(\int_{t_{\bar{k}}}^{t_f} \Psi_{\bar{k}} \nu_{\bar{k}}(t) dt \right) + \\ & \int_{t_{\bar{k}}}^{t_f} \nu_{\bar{k}}^T(t) \aleph_{i\bar{k}}^u \nu_{\bar{k}}(t) dt \leq \\ & 4t_f \varrho_{\bar{k}}^2 \|\aleph_{i\bar{k}}^u\| (\Psi_{\max}^{\bar{k}})^2 + \varpi_{\bar{k}} \|\aleph_{i\bar{k}}^u\|, \quad (21a) \\ & \frac{1}{2} \int_{t_{\bar{k}}}^{t_f} \|\Delta \hat{\omega}_{i\bar{k}}\|_{R_i^\omega}^2 dt = \\ & \frac{1}{2} \int_{t_{\bar{k}}}^{t_f} \left\| -2 \int_{t_{\bar{k}}}^{t_f} \Psi_{\bar{k}} \nu_{\bar{k}}(t) dt + \nu_{\bar{k}}(t) \right\|_{R_i^\omega}^2 dt \leq \end{aligned}$$

$$4t_f \left(\int_{t_k}^{t_f} \Psi_k \nu_k(t) dt \right)^T \mathbb{N}_{i\bar{k}}^\omega \left(\int_{t_k}^{t_f} \Psi_k \nu_k(t) dt \right) + \int_{t_k}^{t_f} \nu_k^T(t) \mathbb{N}_{i\bar{k}}^\omega \nu_k(t) dt \leq 4t_f \varrho_k^2 \|\mathbb{N}_{i\bar{k}}^\omega\| (\Psi_{\max}^k)^2 + \varpi_k \|\mathbb{N}_{i\bar{k}}^\omega\|, \quad (21b)$$

式中:

$$\int_{t_k}^{t_f} \|\nu_k\|^2 dt \leq \frac{2\hat{V}(t_k) \lambda_{\max}(P_k)}{\bar{\alpha}_k} (1 - e^{-\bar{\alpha}_k t_f}) \triangleq \varpi_k, \int_{t_k}^{t_f} \|\nu_k\| dt \leq \frac{2\sqrt{2\hat{V}(t_k) \lambda_{\max}(P_k)}}{\bar{\alpha}_k} (1 - e^{-\frac{\bar{\alpha}_k t_f}{2}}) \triangleq \varrho_k, \hat{V}_k = \frac{1}{2} \nu_k^T P_k \nu_k, P_k \text{ 为采样时刻 } t_k \in [t_0, t_f] \text{ 内黎卡提方程 } -\Lambda^T P_k - P_k \Lambda = -I_{3N} \text{ 对应的解. 最后, 根据式(21), 可以得出}$$

$$\Delta J_{i\bar{k}}^{\text{SH}} \leq (\Pi_{\max}^k)^2 \|F_i\| + \varrho_k \|\Pi_k \Psi_k F_i\| + \varrho_k^2 \Xi_{i\bar{k}} + \varpi_k \Theta_{i\bar{k}} \triangleq \epsilon_{i\bar{k}}.$$

证毕.

2.9 带有避碰约束的近似LR-NE

在近似LR-NE解(13)中引入额外的避碰约束, 势函数形式如下:

$$\hat{P}_{io}(\check{x}_{io}^k) = (\min\{0, \frac{\|\check{x}_{io}^k\|^2 - A_i^2}{\|\check{x}_{io}^k\|^2 - a_{io}^2}\})^2, \quad (22)$$

式中: $o \in \bar{\mathcal{O}}$ 表示第 i 个参与人在感应范围内的障碍物, $\bar{\mathcal{O}} \triangleq \mathcal{O} \cup \bar{\mathcal{V}}$; A_i 表示第 i 个参与人的感应半径; $a_{io} = a_i + a_o$ 表示第 i 个参与人与障碍物 o 之间的安全距离, a_i, a_o 分别表示第 i 个参与人与障碍物 o 的安全半径; $\check{x}_{io}^k = x_{i\bar{k}} - x_o$ 表示第 i 个参与人与障碍物 o 之间的距离, x_o 表示参与人感应到障碍物 o 边缘最近的位置.

对上式(22)关于 \check{x}_{io}^k 求偏导, 可以推导出

1) 当 $a_{io} \leq \|\check{x}_{io}^k\| \leq A_i$ 时,

$$\frac{\partial \hat{P}_{io}(\check{x}_{io}^k)}{\partial \check{x}_{io}^k} = \frac{4(A_i^2 - a_{io}^2)(\|\check{x}_{io}^k\|^2 - A_i^2)}{((\check{x}_{io}^k)^{-1})^T (\|\check{x}_{io}^k\|^2 - a_{io}^2)}; \quad (23)$$

2) 当 $\|\check{x}_{io}^k\| \geq A_i$ 时, $\frac{\partial \hat{P}_{io}(\check{x}_{io}^k)}{\partial \check{x}_{io}^k} = 0$.

进一步, 基于式(13)与式(23)可以得出带有避碰约束的近似LR-NE, 如下:

$$\hat{u}_{i\bar{k}}^* = -\Upsilon_i^u F_i(\hat{x}_k - \bar{x}^*) - \sum_{o=1}^{|\bar{\mathcal{O}}|} \frac{\partial \hat{P}_{io}^T(\check{x}_{io}^k)}{\partial \check{x}_{io}^k}. \quad (24)$$

2.10 MASs的稳定性分析

定理 4 考虑由 N ($N \geq 2$) 个参与人组成的博弈系统, 其中每个参与人受式(1)–(4)以及网络拓扑结构 $G(\mathcal{V}, \varepsilon)$ 约束, 个体博弈成本满足式(9). 在采样时刻 $t_k \in [t_0, t_f]$ 内, 第 i 个参与人的近似LR-NE满足式(24), 则MASs(3)可以无碰的实现期望的编队目标.

证 令候选李雅普诺夫函数为

$$\bar{V} = \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \frac{1}{4} (\tilde{x}_{i\bar{k}} - \tilde{x}_{o\bar{k}})^T (\tilde{x}_{i\bar{k}} - \tilde{x}_{o\bar{k}}) + \frac{1}{2} \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \hat{P}_{io}(\check{x}_{io}^k), \quad (25)$$

式中 $\tilde{x}_{i\bar{k}}$ 与 $\tilde{x}_{o\bar{k}}$ 分别表示第 i 个参与人与第 o 个参与人的状态误差.

当第 o 个参与人为第 i 个参与人感应范围内的智能体时, 对式(25)沿系统轨迹求偏导, 可以得出

$$\dot{\bar{V}} = \frac{1}{2} \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \left[\frac{1}{2} (\tilde{x}_{i\bar{k}} - \tilde{x}_{o\bar{k}})^T (\dot{\tilde{x}}_{i\bar{k}} - \dot{\tilde{x}}_{o\bar{k}}) \right] + \frac{1}{2} \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \left(\frac{\partial \hat{P}_{io}}{\partial x_{i\bar{k}}} \dot{x}_{i\bar{k}} + \frac{\partial \hat{P}_{io}}{\partial x_{o\bar{k}}} \dot{x}_{o\bar{k}} \right). \quad (26)$$

根据式(23)可以得出

$$\frac{\partial \hat{P}_{io}}{\partial x_{i\bar{k}}} = -\frac{\partial \hat{P}_{io}}{\partial x_{o\bar{k}}}. \quad (27)$$

由于式(5)成立, 因此

$$\dot{\bar{V}} \leq 0. \quad (28)$$

当第 o 个参与人为第 i 个参与人感应范围内的静态障碍物(速度为0的静态参与人)时, 对式(25)沿系统轨迹求偏导, 可以得出

$$\dot{\bar{V}} = \frac{1}{2} \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \left[\frac{1}{2} (\tilde{x}_{i\bar{k}} - \tilde{x}_{o\bar{k}})^T (\dot{\tilde{x}}_{i\bar{k}} - \dot{\tilde{x}}_{o\bar{k}}) \right] + \frac{1}{2} \sum_{i=1}^N \sum_{o=1}^{|\bar{\mathcal{O}}|} \left(\frac{\partial \hat{P}_{io}}{\partial x_{i\bar{k}}} \dot{x}_{i\bar{k}} \right). \quad (29)$$

由于 $x_i(t_f) \in \mathcal{M}_{i\bar{k}}$, 可以得出, 当 $t \rightarrow t_f$ 时, $\frac{\partial \hat{P}_{io}}{\partial x_{i\bar{k}}} = 0$, 类似地, 可以得出 $\dot{\bar{V}} \leq 0$. 综上所述, MASs可以无碰的实现期望的编队目标. 证毕.

3 仿真实验

本小节考虑两组数值仿真以验证所提R-DRDG方法的有效性和优越性. 第1组与带有障碍物约束但成本函数中没有考虑干扰的RDDG方法(即文献[12]的RDDG方法中引入障碍物约束)对比, 旨在验证R-DRDG的有效性(鲁棒性). 第2组与文献[11]中的DRDG方法比较, 旨在验证R-DRDG的优越性(安全性), 两组案例中的外部干扰为 $\omega_i = [\sin t \ \sin t \ \sin t]^T$, 采样时间均为0.1 s.

3.1 鲁棒性

在该组仿真中, 采用所提出的R-DRDG方法与RDDG方法在具有障碍物的环境下进行仿真对比. 其中, 以3个智能体为例, 通过对比期望的任务目标是否完成、是否避开障碍物、各个智能体与障碍物之间的距离来说明所提方法的有效性. 各个智能体的动态方

程表达式如下:

$$\dot{x}_i = u_i + \omega_i, \forall i \in N_{1:3}. \quad (30)$$

智能体之间的通信拓扑结构如图2所示.

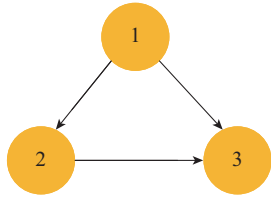


图2 3个智能体的通信拓扑图

Fig. 2 The communication topology of three agents

仿真中使用的相关参数在表1中列出.

表1 3个智能体的仿真参数值

Table 1 Simulation parameter values of three agents

参数	数值
智能体安全半径 a_i	0.5 m
静态障碍物安全半径 a_o	1 m
智能体感应半径 A_i	10 m
终端成本对应权重 F_i	$25I_3$
控制成本对应权重 R_i^u	$7I_3$
干扰成本对应权重 R_i^w	$5I_3$
博弈终端时间 t_f	95 s
O1, O2位置	$[10 \ 10 \ 5]^T$ (m), $[60 \ 12 \ 5]^T$ (m)
相对位置 d_{21}^d, d_{31}^d	$[20 \ 0 \ 0]^T$ (m), $[10 \ 10 \ 0]^T$ (m)
相对位置 d_{32}^d	$[-10 \ 10 \ 0]^T$ (m)

在RDDG方法与R-DRDG方法下规划的编队轨迹如图3所示. 其中: 各个智能体分别用A1, A2与A3表示, 椭球形表示环境中的障碍物, 分别用O1, O2表示障碍物1与障碍物2, O2表示未知障碍物. 通过观察图3可知, 两种方法均使各个智能体收敛到了期望的编队队形.

然而, RDDG方法在躲避O2时, 使A2与O2之间的距离小于安全距离, 如图4所示. 这是由于RDDG方法的成本函数中未考虑干扰成本, 不能消解外界环境对智能体运行造成的影响. 然而, R-DRDG方法是在最大化干扰成本的情况下求解最优鲁棒控制策略. 因此, R-DRDG方法使智能体与障碍物之间的距离均大于安全距离.

进一步, 通过各个智能体之间的状态误差衡量R-DRDG方法与RDDG方法是否能够使智能体实现期望的编队队形, 如图5所示. 通过观察图5可知, 两种方法几乎均能使MASs近似收敛到0, 说明所设计的终端状态估计器是有效的. 然而, 根据局部放大图可知, RDDG方法的编队收敛状态误差比R-DRDG方法的收敛状态误差大, 这是由于RDDG方法的代价函数中没有考虑干扰的影响, 而R-DRDG方法通过消解干扰对MASs的影响, 使得编队可以收敛到期望的状态.

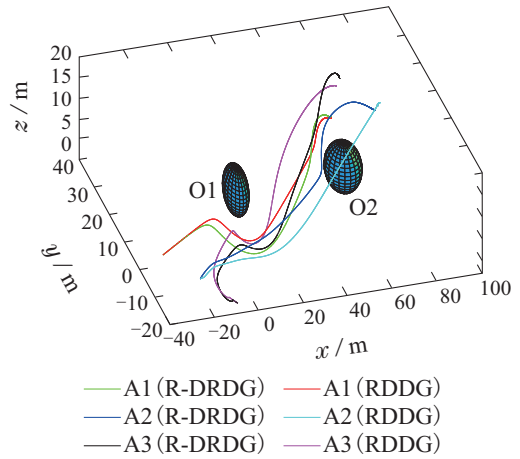


图3 智能体在两种方法下的轨迹

Fig. 3 The trajectories of the agents with two methods

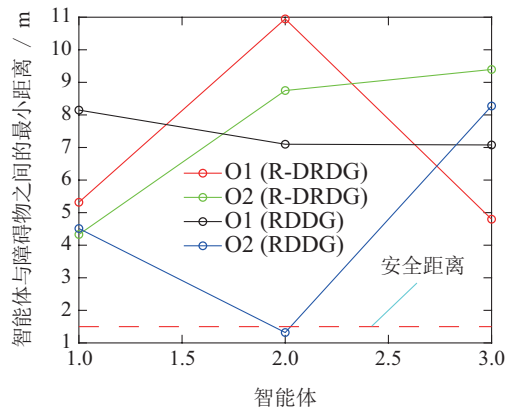


图4 智能体与障碍物之间的最小距离

Fig. 4 The minimum distance between agents and obstacles

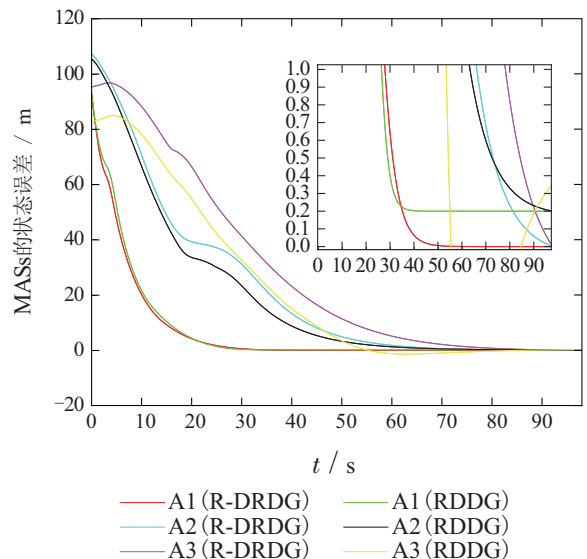


图5 多智能体系统的状态误差

Fig. 5 The state error of MASs

3.2 安全性

在该仿真中, 采用所提出的R-DRDG方法与文献[11]中的DRDG方法在未知障碍物环境下进行仿真对

比. 其中, 以8个智能体为例, 通过对比期望的任务目标是否完成、是否避开障碍物、各个智能体与障碍物之间的距离来说明所提方法的避碰性能. 各个智能体的动态方程与式(30)相同, 此外, 智能体的安全距离、感应半径、障碍物的安全距离以及博弈终端时间与仿真案例1的参数设置相同. MASs的通信拓扑结构如图6所示.

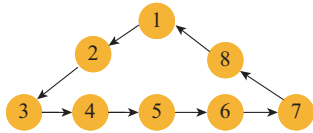


图6 8个智能体的通信拓扑图

Fig. 6 The communication topology of eight agents

博弈成本对应的相关权重分别是 $F_i = 25I_3$; $R_i^u = 5I_3$; $R_i^\omega = 5I_3$. 分布式估计器增益 $s_i = 25$, 仿真环境中设置的障碍物位置以及智能体之间的期望相对位置在表2中列出.

表2 8个智能体的仿真参数值

Table 2 Simulation parameter values of eight agents

参数	数值
O1, O2位置/m	$[15 \ 18 \ -16]^T, [29 \ 16 \ -20]^T$
O3, O4位置/m	$[11 \ -10 \ -30]^T, [20 \ 43 \ -13]^T$
O5, O6位置/m	$[-5 \ -18 \ -30]^T, [-10 \ 40 \ 17]^T$
相对位置 d_{21}^d, d_{32}^d /m	$[-10 \ -10 \ 0]^T, [-10 \ -10 \ 0]^T$
相对位置 d_{43}^d, d_{54}^d /m	$[10 \ 0 \ 0]^T, [10 \ 0 \ 0]^T$
相对位置 d_{65}^d, d_{76}^d /m	$[10 \ 0 \ 0]^T, [10 \ 0 \ 0]^T$
A7-A8相对位置 d_{87}^d /m	$[-10 \ 10 \ 0]^T$

各个智能体使用DRDG方法规划的编队轨迹如图7所示. 其中, 环境中的障碍物、智能体图标与仿真案例一表示一致, O1表示未知障碍物. 根据图7可知, A8使用DRDG方法规划编队轨迹时, 由于反馈策略的局部本质, 仅利用当前状态与邻居参与者博弈, A8在感应到O1时, 已经陷入死锁区域(与O1发生碰撞, 陷入死区, 无法前进), 导致A8停滞在O1附近. 如图7中的局部放大图, A8进入死锁状态.

各个智能体使用R-DRDG方法规划的编队轨迹如图8所示. R-DRDG方法由于在滚动时域窗口内, 可以基于未来一段时间的状态信息与邻居进行博弈, 提前避开未来几个时间步即将出现的碰撞风险. 当智能体在运行过程中, 可以在未来时间步内提前感应到未知障碍物的出现, 进一步提前做好避碰准备, 避免由于仅利用当前状态及时避碰而导致陷入死锁的风险, 最终可以使各个智能体成功的避开未知障碍物, 如图8中的局部放大图, A8躲避了潜在的死锁风险.

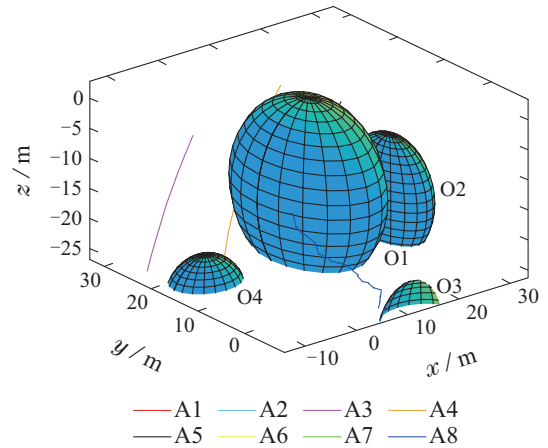
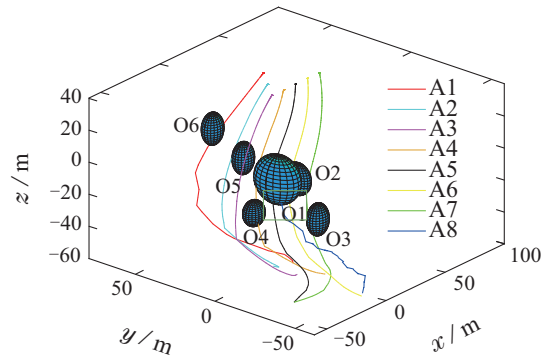


图7 智能体在DRDG方法下的轨迹

Fig. 7 The trajectories of agents with the DRDG method

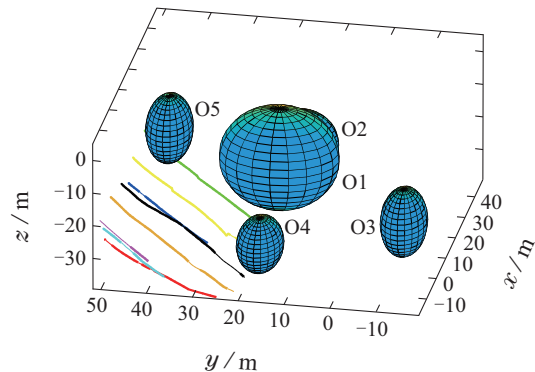
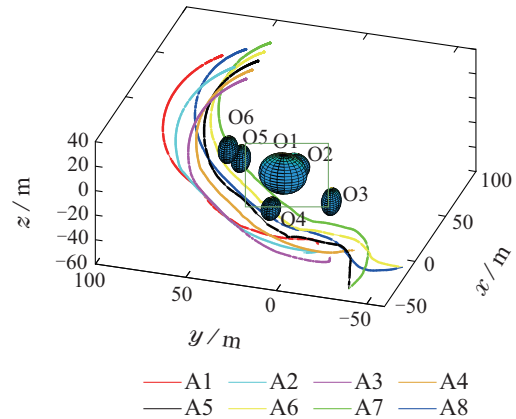


图8 智能体在R-DRDG方法下的轨迹

Fig. 8 The trajectories of agents with the R-DRDG method

图9表示使用DRDG方法时, 各个智能体与障碍物

之间的最小距离. 根据图9可知, 由于DRDG方法中反馈策略仅根据当前获取的状态信息求解鲁棒NE策略, 具有局部本质, 使A8陷入死锁区域, 进一步与O1之间的距离小于安全距离.

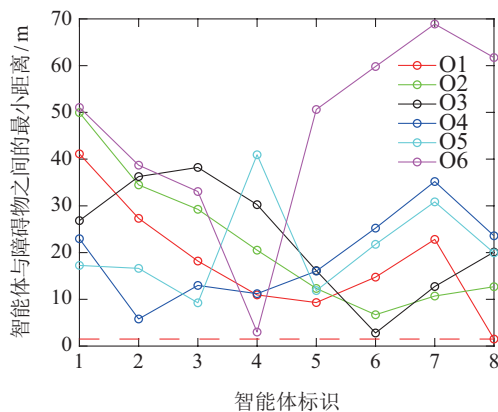


图9 智能体在DRDG方法下与障碍物的最小距离

Fig. 9 The minimum distance between agents and obstacles with the DRDG method

进一步通过分析图10可知, 在R-DRDG方法下, 各个智能体与障碍物之间的最小距离均大于安全距离, 说明通过在滚动窗口内重复博弈, 可以利用未来状态信息提前躲避潜在的碰撞风险, 进一步提升智能体运行的安全性能.

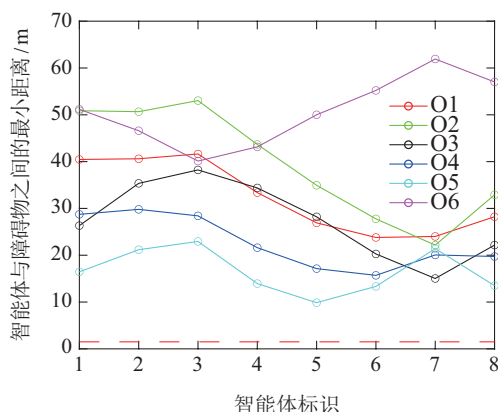


图10 智能体在R-DRDG方法下与障碍物的最小距离

Fig. 10 The minimum distance between agents and obstacles with the R-DRDG method

4 结论

本文提出了一种滚动分布式鲁棒微分博弈方法, 用于解决MASs在未知障碍物环境下, 受通信以及干扰约束的编队控制问题. 通过优化在最坏情况下的性能指标, 实现了对外部干扰的鲁棒性. 此外, 基于分布式终端状态估计器, 获得了仅依赖邻居状态信息的近似LR-NE解. 随后, 通过严格的数学证明推导了近似LR-NE的收敛性. 进一步在近似LR-NE解中引入额外的避碰惩罚项, 以保证智能体的安全运行. 最后, 在收

缩时域窗口内利用对未知状态信息的预测, 减少了死锁现象的发生, 提升了智能体的运行过程中的安全性能. 通过仿真表明, R-DRDG方法的鲁棒性与安全性能明显优于目前流行的RDDG, DRDG方法. 在未来的研究工作中, 一方面考虑如何结合多层网络设计适应大规模MASs的分布式状态估计器; 另一方面考虑如何设计带有干扰约束的事件触发机制, 解决滚动鲁棒博弈模型中存在实时性不佳的问题.

参考文献:

- [1] HUANG J, FANG H, DOU L, et al. An overview of distributed high-order multi-agent coordination. *IEEE/CAA Journal of Automatica Sinica*, 2014, 1(1): 1–9.
- [2] SHI Q, LI T, LI J, et al. Adaptive leader-following formation control with collision avoidance for a class of second-order nonlinear multi-agent systems. *Neurocomputing*, 2019, 350(1): 282–290.
- [3] YAO P, WEI Y, ZHAO Z. Null-space-based modulated reference trajectory generator for multi-robots formation in obstacle environment. *ISA Transactions*, 2022, 123(1): 168–178.
- [4] NA Qiantai, GAO Fei, WENG Zhi, et al. Distributed cooperative control of incomplete-constrained multi-agent systems based on barrier control functions. *Control Theory & Applications*, 2022, 39(4): 663–670.
(娜茜泰, 高飞, 翁智, 等. 非完整约束多智能体系统基于屏障控制函数的分布式协同控制. *控制理论与应用*, 2022, 39(4): 663–670.)
- [5] CAPPELLO D, GARCIN S, MAO Z, et al. A hybrid controller for multi-agent collision avoidance via a differential game formulation. *IEEE Transactions on Control Systems Technology*, 2020, 29(4): 1750–1757.
- [6] GU D. A differential game approach to formation control. *IEEE Transactions on Control Systems Technology*, 2007, 16(1): 85–93.
- [7] LI Y, HU X. A differential game approach to intrinsic formation control. *Automatica*, 2022, 136: 110077.
- [8] GARCIA E, CASBEER D W, VON MOLL A, et al. Multiple pursuer multiple evader differential games. *IEEE Transactions on Automatic Control*, 2020, 66(5): 2345–2350.
- [9] SINGH S K, REDDY P V, VUNDURTHY B. Study of multiple target defense differential games using receding horizon-based switching strategies. *IEEE Transactions on Control Systems Technology*, 2021, 30(4): 1403–1419.
- [10] JIAO Q, MODARES H, XU S, et al. Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control. *Automatica*, 2016, 69(1): 24–34.
- [11] XUE W, ZHAN S, WU Z, et al. Distributed multi-agent collision avoidance using robust differential game. *ISA Transactions*, 2023, 134(1): 95–107.
- [12] LIN W, LI C, QU Z, et al. Distributed formation control with open-loop nash strategy. *Automatica*, 2019, 106(1): 266–273.

作者简介:

薛文艳 讲师, 博士, 目前研究方向为博弈论与多智能体协调控制, E-mail: wyxuedu@163.com;

黄捷 教授, 博士, 目前研究方向为多智能体系统协同控制、工业互联网理论与技术, E-mail: jie.huang@fzu.edu.cn.