具有地面覆盖保证的有限状态分布式策略控制

Hossein Rastgoftar

arXiv:2601.02109v1 [eess.SY] 5 Jan 2026

摘要—**我们提出一种用于多智能体地面覆盖的有限状态分布式决策与控制框架。该方法将问题分解为两个相互耦合的组成部分：(i) 根据智能体的参考构型设计深度神经网络（DNN）的结构，以及 (ii) 基于策略的分布式覆盖控制。智能体分为锚点和跟随者，由此形成一种通用且可扩展的通信架构：每个跟随者恰好与来自前一层的三个入邻居交互，构成封闭的三角形通信结构。深度神经网络训练权重隐式编码了智能体团队的空间构型，从而提供环境目标集的几何表示。在此架构中，我们建立了一种计算高效的分布式马尔可夫决策过程（MDP）；其组成部分均为时不变，唯有时变代价函数例外，该函数由每个智能体的通信三角形所包含目标集的质心偏差定义。通过引入任意输出可控性（AOC）的概念，我们假设每个智能体均具有AOC，并建立了去中心化收敛性，使其收敛到能够最优表示环境目标的期望构型。**

I. 引言

多智能体地面覆盖是分布式控制中的一个基本问题，应用于环境监测、监视和分布式感知。经典且广泛采用的一种方法是基于沃罗诺伊图的覆盖控制，智能体通过迭代地向各自沃罗诺伊单元的质心移动，以分散式方式优化空间覆盖目标。这一范式具有很强的几何可解释性和收敛保证，已得到广泛研究与拓展，包括密度加权覆盖、受限环境以及事件触发式实现 [7], [14], [29]。

尽管具有这些优势，基于沃罗诺伊图的方法通常依赖连续时间动力学、频繁的邻居重计算以及显式几何构造，这些因素限制了其在通信受限条件下的可扩展性，也使其难以与离散决策和学习机制集成。这些局限促使人们开发基于策略和学习增强的分散式框架，包括有限状态和马尔可夫决策过程表述，以应对多智能体协同中的不确定性和可扩展性问题 [22], [28]。尽管这类方法更加灵活，但往往缺乏明确机制，无法在分散式策略中编码编队几何结构，也无法施加具有可解释性和结构化的信息流，同时保证收敛性。

本文通过引入一种结构化的、基于策略的框架来应对这些挑战。该框架紧密耦合智能体间通信、决策与物理演化，从而在保留几何意义和解析可处理性的同时，实现可扩展的分布式覆盖。

A. 相关工作

文献 [11], 报告了多智能体覆盖的扩散式收敛与稳定性结果；文献 [5]研究了采用局部密度反馈和平均场近似的分布式覆盖。文献 [3]探讨了领导者—跟随者覆盖策略，包括明确区分协调目标与覆盖目标。文献 [10], [26],研究了自适应分布式覆盖方法，文献 [16]则给出了多尺度连续时间收敛分析。文献 [12], [25]介绍了相关方法在人本感知和区域覆盖规划中的应用。大量研究采用基于沃罗诺伊图的覆盖控制 [1], [4], [18], [19],，通常在运动学或单积分器智能体抽象下，通过基于李雅普诺夫的方法证明收敛性。文献 [4], 考虑了障碍物、故障和领导者—跟随者结构等方面的扩展；文献 [21]则报告了复杂城市环境中的实验对比结果。异构智能体团队的覆盖控制也日益受到关注。文献 [24] 的作者提出了一种异构覆盖控制框架，通过将智能体特定的密度函数纳入位置代价函数，以编码不同的感知能力；他们据此推导出一种分布式梯度下降控制器，并加入额外的边界项，以确保收敛至异构覆盖目标的临界点。实验结果表明，该方法优于异构版 Lloyd 方法。文献 [2]提出了一种针对异构圆盘形机器人的基于沃罗诺伊图的覆盖控制方法，该方法利用幂图和受约束的质心运动，确保无碰撞地收敛至局部最优感知配置。Sadeghi 和 Smith 针对多种事件类型的异构机器人覆盖控制，构建了特定于事件的沃罗诺伊划分框架，并推导出分布式算法；该算法在连续和离散环境中均可证明收敛至局部最优感知配置 [23]。文献 [15], 提出了一种面向最高速度各异机器人的覆盖控制框架，通过乘法加权沃罗诺伊图构造时间代价，并推导出基于梯度的控制器，从而得到时间最优的覆盖配置。较新的研究则关注多资源目标和持续监视目标 [6], [13]。

基于学习的方法将多智能体覆盖构建为采用强化学习的决策过程

Hossein Rastgoftar 任职于美国亚利桑那州图森市 85719 的航空航天与机械工程系hrastgoftar@arizona.edu

和马尔可夫决策模型 [8], [9], [17], [20], [27]。尽管这些方法具有可扩展性和适应性，但通常依赖非结构化通信、大规模或连续状态空间以及基于梯度的优化，因此限制了其可解释性和收敛性分析。

为阐明本文与现有基于学习的覆盖方法之间的区别，我们强调，本文并未将多智能体覆盖视为一般性的强化学习或函数逼近问题。相反，智能体间通信通过由参考构型所诱导的层级式锚点—跟随者结构进行明确设计，从而形成单向前馈信息流。这种结构使多智能体系统本身可被解释为动力学神经网络，其中神经元对应于物理智能体，其激活由智能体动力学而非代数映射决定。学习通过仅前向传播的局部更新完成，不涉及梯度反向传播或集中式评论家；每个智能体则求解一个在其通信三角形内以几何方式定义的有限时不变局部马尔可夫决策过程。这些特征从根本上使所提出的框架区别于现有的强化学习和基于马尔可夫决策过程的覆盖方法。

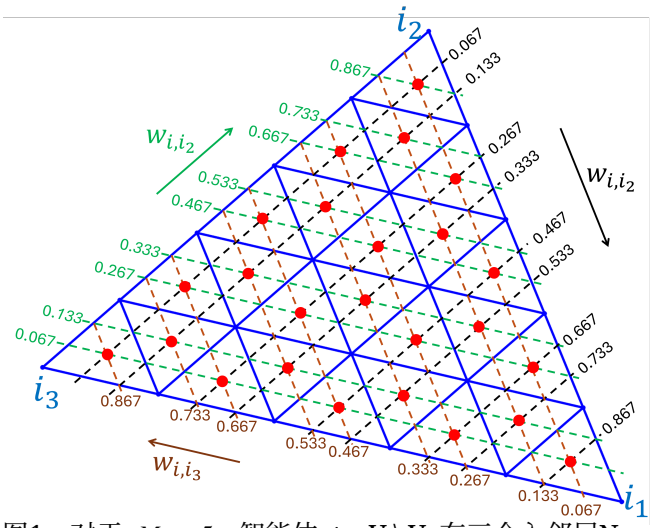


图1：对于 $M_i = 5$，智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 有三个入邻居$\mathrm{N}(i) = \{i_1, i_2, i_3\}$时，第一层和第二层通信权重$w_{i,i_1} \in \mathrm{W}_{i,1}$ 和$w_{i,i_2} \in \mathrm{W}_{i,2}$的几何表示。

B. 贡献

本文提出了一种基于策略的去中心化框架，用于覆盖未知地面目标；该框架可扩展至任意规模的团队，并且不依赖于单个智能体的动力学。其核心思想是将多智能体覆盖重新诠释为一种结构化动力系统，在该系统中，通信、决策和物理演化内在耦合。通过依据参考构型组织智能体间通信，所提出的方法形成一种层级式前馈协调架构，可被解释为动力学深度神经网络，同时仍保持完全去中心化。

主要贡献总结如下：

- **结构化通信与动力学深度神经网络表示：**本文提出一种分层锚点—跟随者通信架构，该架构形成单向前馈信息流，使多智能体系统可被解释为动力学神经网络，其中神经元对应于物理智能体。

- **基于局部可观测性的仅前向学习：**通信权重仅通过前向局部更新进行学习，无需梯度反向传播、集中式评论家或全局信息。

- **通过局部马尔可夫决策过程进行分散式策略学习：**每个跟随者智能体通过求解在其通信三角形内以几何方式定义的有限、时不变局部马尔可夫决策过程，独立学习转移策略。

- **通过任意输出可控性实现与动力学无关的覆盖：**任意输出可控性这一概念将策略学习与特定智能体动力学解耦，使其能够统一应用于具有非线性、欠驱动或高阶动力学的异构团队。

C. 概述

本文组织如下：第II节阐述问题。第III节提出一种算法方法，根据智能体团队的参考构型构建深度神经网络。第IV节将深度神经网络权重训练定义为一个马尔可夫决策过程并加以介绍。第V节证明所提出的基于策略的分散式覆盖方案的稳定性与收敛性。第VI节给出仿真结果，最后在第VII节总结全文。

II. 问题陈述

我们考虑一支由 N 个智能体组成的团队，编号为

$$\mathcal{V} = \{1, \ldots, N\},$$

其任务是为有限集合 D 中的地面目标提供空中覆盖。根据参考构型，智能体分为边界 或内部 智能体，并通过德劳内邻居网络（DNN）组织其交互，以实现可扩展的覆盖。低层控制动力学被抽象化，并假设每个智能体均满足以下输出可达性。

定义 1（任意输出可控性（AOC））。 设智能体 $i \in \mathrm{V}$ 的描述如下：

$$\begin{cases} \mathbf{x}_i[t+1] = \mathbf{f}_i(\mathbf{x}_i[t], \mathbf{u}_i[t]), \\ \mathbf{r}_i[t] = \mathbf{h}_i(\mathbf{x}_i[t]), \end{cases} \tag{1}$$

其中 $\mathbf{x}_i$、$\mathbf{u}_i$ 和 $\mathbf{r}_i$ 分别表示状态、输入和输出。若对于任意初始状态 $\mathbf{x}_i[t] \in \mathrm{X}_i$，都存在一个容许输入序列 $\mathbf{u}_i(\cdot)$ 和有限时间 $T_i(\mathbf{x}_i[t]) < \infty$，使得以下条件成立，则称智能体 i 具有任意输出可控性：

$$\mathbf{r}_i[t + T_i(\mathbf{x}_i[t])] \in \mathcal{P}_i,$$

其中，$\mathrm{X}_i \subset \mathbb{R}^2$ 和 $\mathrm{P}_i \subset \mathbb{R}^2$ 是紧集。

假设 1。 所有智能体采用统一的时间离散化，且离散化步长足够大，使得输出可达时间满足

$$T_i(\mathbf{x}_i[t]) = 1,$$

对于所有 $i \in \mathrm{V}$ 以及所有初始状态 $\mathbf{x}_i[t] \in \mathrm{X}_i$。因此，

$$\mathbf{r}_i[t+1] \in \mathcal{P}_i,$$

对于任意可容许初始条件均成立。

本文旨在设计一种分布式框架，以实现结构化的智能体交互和对分布式目标的自适应覆盖。具体而言，我们将解决以下问题。

问题*1*（深度神经网络结构构建）。

给定一个参考构型，确定性算法根据智能体的初始空间分布唯一确定一种深度神经网络通信架构。智能体集合 V 被划分为 $M+1$ 互不相交的子集

$$\{\mathcal{V}_l\}_{l=0}^{M}, \qquad \bigcup_{l=0}^{M} \mathcal{V}_l = \mathcal{V},$$

其中，V_0 由锚节点组成，并且每个智能体 $i \in \mathrm{V}_l, l \geq 1$ 均恰有三个入邻居。

问题*2*（通过离散深度神经网络权重实现分布式覆盖）。

设计一种分散式控制与学习机制，使智能体团队能够对分布式目标集 D 实现高水平覆盖。对于每个智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0,$，深度神经网络训练权重被限制在有限离散集中

$$\mathcal{W}_i = \left\{ \frac{3a-b}{3M_i} : a = 1, \cdots, M_i,\ b = 1, 2 \right\} \tag{2}$$

其中 $M_i \in \mathbb{N}$ 决定离散化分辨率。这些集合 W_i 由 (0,1) 中均匀分布的值构成，从而确保训练权重严格为正且有界。对于每个智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0,$，令 $\mathrm{N}(i)$ 表示其入邻居集合。智能体 i 与邻居 $j \in \mathrm{N}(i)$之间的通信权重记为 $w_{i,j}$，并满足

$$w_{i,j} \in \mathcal{W}_i, \qquad \forall i \in \mathcal{V} \setminus \mathcal{V}_0,\ \forall j \in \mathcal{N}(i), \tag{3a}$$

$$\sum_{j \in \mathcal{N}(i)} w_{i,j} = 1, \qquad \forall i \in \mathcal{V} \setminus \mathcal{V}_0. \tag{3b}$$

为便于说明，图1展示了 $M_i = 5$的通信权重的几何表示，其中智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$与三个入邻居 $\mathrm{N}(i) = \{i_1, i_2, i_3\}$ 进行交互。相应的离散权重集

$$\mathcal{W}_i = \{0.067, 0.133, 0.267, 0.333, 0.467, 0.533, 0.667, 0.733, 0.867\}$$

由式（2）得到。问题2将覆盖问题公式化为一个分布式马尔可夫决策过程，其中状态空间、动作空间、状态转移模型和折扣因子均为时不变，而代价函数随时间变化，用于反映覆盖性能。详细的马尔可夫决策过程公式化及深度神经网络权重训练过程见第IV节。

III. 覆盖深度神经网络的结构构建

基于参考构型，将智能体集合 V 划分为由 M 索引的互不相交的组 $M+1 = \{0, 1, \ldots, M\}$，从而得到 DNN通信架构。这引出如下分解:

$$\mathcal{V} = \bigcup_{l \in \mathcal{M}} \mathcal{V}_l, \qquad \mathcal{V}_l \cap \mathcal{V}_h = \emptyset,\ l \neq h,$$

满足 $\mathrm{V}_l \subset \mathrm{V}$，且对所有 $l \in \mathrm{M}$，基数为 $|\mathrm{V}_l| = N_l$。定义累积索引

$$P_l = \begin{cases} \sum_{h=0}^{l} N_h, & l \in \mathcal{M} \setminus \{0\}, \\ 0, & l = 0, \end{cases}$$

并将智能体依次编号为 $\{b_1$、.....、$b_N\}$。然后

$$\mathcal{V}_l = \{b_{P_{l-1}+1}, \ldots, b_{P_l}\}. \tag{4}$$

为定义智能体间通信，引入嵌套集合

$$\mathcal{L}_l = \begin{cases} \mathcal{V}_l, & l \in \{0, M\}, \\ \mathcal{V}_l \cup \mathcal{L}_{l-1}, & \text{otherwise}, \end{cases} \qquad \forall l \in \mathcal{M}. \tag{5}$$

令 $\mathrm{I}(i,l) \subseteq \mathrm{L}_{l-1}$表示第 $l-1$ 层中与神经元 $i \in \mathrm{L}_l$相连的神经元集合。深度神经网络架构由智能体的初始队形通过算法1中的算法流程导出，该流程构建了一个可用深度神经网络表示的有向图 G(V,E)。具体而言，G(V,E)确定了: (i) 深度神经网络的层数（$M+1$），(ii) 将 V 划分为若干子集 $\mathrm{V}_0, \ldots\ldots, \mathrm{V}_M$，以及 (iii) 层间神经元连接关系。

给定 $\mathrm{E} \subset \mathrm{V} \times \mathrm{V}$，智能体 $i \in \mathrm{V}$ 的入邻居集合定义为

$$\mathcal{N}(i) := \{ j \in \mathcal{V} \mid (j,i) \in \mathcal{E} \}. \tag{6}$$

然后，对于每个层 $l \in \mathrm{M} \setminus \{0\}$，神经元 $i \in \mathrm{L}_l$ 的互连集合 $\mathrm{I}(i, l)$ 定义为

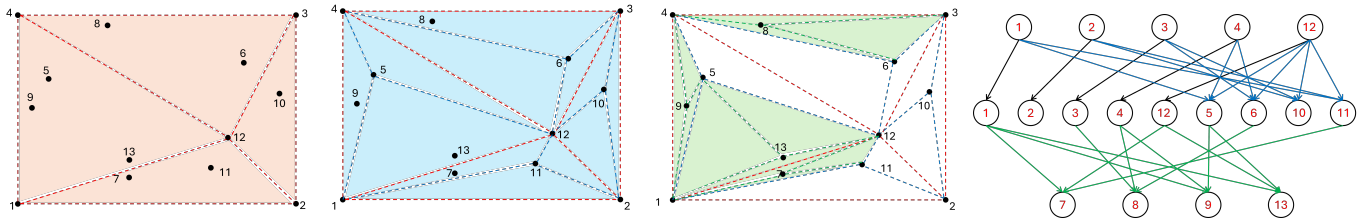
$$\mathcal{I}(i,l) = \begin{cases} \mathcal{N}(i), & i \in \mathcal{V}_l = \mathcal{L}_l \setminus \mathcal{L}_{l-1}, \\ \{i\}, & i \in \mathcal{L}_l \setminus \mathcal{V}_l, \end{cases} \qquad l \in \mathcal{M} \setminus \{0\}. \tag{7}$$

注释 1。 算法1适用于 $\mathbb{R}^n$ 中的分散式多智能体系统；此处考虑的地面覆盖场景对应于 $n = 2$。

对于每个智能体 $i \in \mathrm{V}$，本文始终使用以下与位置相关的量:

- $\mathbf{a}_i$: 智能体 i 在初始（参考）配置中的参考位置。
- $\mathbf{r}_i[t]$: 智能体 i 在离散时间 t 的实际位置，由其控制系统的输出给出。
- $\mathbf{c}_i[t]$: 智能体 i 在时间 t 输入其控制系统的参考输入；对于 $i \in \mathrm{V}_0$，$\mathbf{c}_i[t]$ 为常量；而对于 $i \in \mathrm{V} \setminus \mathrm{V}_0$，其定义为其入邻居智能体实际位置的加权平均。
- $\mathbf{p}_i$: 智能体的期望位置 i；$\mathbf{p}_i$ 在 $i \in \mathrm{V}_0$ 中已知，在 $i \in \mathrm{V} \setminus \mathrm{V}_0$中未知。

此外，对于所有 $i \in \mathrm{V} \setminus \mathrm{V}_0,$，参考输入满足$\mathbf{c}_i[t] = \mathbf{p}_i$。



(a) $l=0\in\mathrm{M}$ (b) $l=1\in\mathrm{M}$ (c) $l=2\in\mathrm{M}$ (d) $\mathrm{M}=\{0,\ 1,\ 2\}$

图 2：边界智能体定义的凸包的单元分解，用于指定 DNN 层互连。

A. 步骤 *1*：智能体分类

智能体集合 V 被分解为 $\mathrm{V}=\mathrm{V}_B\cup\mathrm{V}_I$，其中 V_B 和 V_I 不相交。集合 $\mathrm{V}_B=\{b_1,\ \ldots\ldots,\ b_{N_B}\}$ 由位于凸包顶点处的边界智能体组成，该凸包包围 V_I 中的内部智能体。由 V_B 定义的多胞体称为前导多胞体。给定边界集合 V_B，根据以下两种准则之一确定核心智能体：(i) 总距离最小的内部智能体，

$$b_{N_B+1}=\arg\min_{i\in\mathcal{V}\setminus\mathcal{V}_B}\sum_{j\in\mathcal{V}_B}\|\mathbf{a}_i-\mathbf{a}_j\|, \tag{8}$$

或者（ii）位于目标域 D 中心附近的智能体。

给定集合 V_B 和指定的核心智能体 b_{N_B+1}，我们将集合 V_0 定义为 $\mathrm{V}_0=\mathrm{V}_B\cup\{b_{N_B+1}\}$。根据式（5），可得 $\mathrm{V}_0=\mathrm{L}_0$。前导多胞体可划分为 m_0 个不同的单纯形单元。因此，集合 L_0 可表示为这些单纯形的并集：

$$\mathcal{L}_0=\bigcup_{h=1}^{m_0}\mathcal{R}_{0,h}, \tag{9}$$

其中，$\mathrm{R}_{0,\ h}$ 确定前导多胞体中第h个单纯形单元的顶点。为便于说明，由 $N=13$ 个智能体组成的智能体团队形成了图 2（a）所示的二维编队，其中 $\mathrm{V}_B=\{1,\ \cdots,\ 4\}$ $(N_B=4)$ 表示边界智能体。智能体 $b_5=12\in\mathrm{V}$ 被指定为式（8）中的核心领导者，因此，$\mathrm{V}_0=\mathrm{L}_0=\{1,\ \cdots,\ 4,\ 12\}$表示第一层的智能体。由 L_0 定义的凸包可分解为 $m_0=4$ 个三角形单元，其顶点分别由 $\mathrm{R}_{0,\ 1}=\{1,\ 2,\ 12\}$、$\mathrm{R}_{0,\ 2}=\{2,\ 3,\ 12\}$、$\mathrm{R}_{0,\ 3}=\{3,\ 4,\ 12\}$以及 $\mathrm{R}_{0,\ 4}=\{4,\ 1,\ 12\}$定义。

B. 步骤 *2*：扩展与结构化

集合 V 可表示为 $\mathrm{V}=\mathrm{L}_{l-1}\cup\mathrm{L}_{l-1}$，对于每个$l\in\mathrm{M}\setminus\{0\}$，其中 $\mathrm{L}_{l-1}=\mathrm{V}\setminus\mathrm{L}_{l-1}$ 表示不属于 L_{l-1}的智能体。注意，若 $\mathrm{L}_{l-1}\neq\emptyset$，则 $\mathrm{V}_l\subset\mathrm{L}_{l-1}$。此外，$\mathrm{L}_{l-1}$ 由 m_{l-1} 个互不相同的单纯形组成，它们覆盖了 L_{l-1}所包含的区域。因此，L_{l-1} 可表示为：

$$\mathcal{L}_{l-1}=\bigcup_{h=1}^{m_{l-1}}\mathcal{R}_{l-1,h},\qquad l\in\mathcal{M}\setminus\{0\} \tag{10}$$

其中，$\mathrm{R}_{l-1,1}$ 至 $\mathrm{R}_{l-1,m_{l-1}}$ 是不同单纯形单元的顶点，这些单元覆盖了 L_{l-1}所包含的区域。给定一组 $\mathrm{R}_{l-1,h}$，对于每个 $h=1,\ldots,m_{l-1}$，我们将由 $\mathrm{R}_{l-1,h}$中的元素构成的凸包记为 CONV($\mathrm{R}_{l-1,h}$)。我们还将 $\mathrm{H}_{l-1,h}\subset\mathrm{L}_{l-1}$ 定义为位于该凸包内的所有节点的集合，即包含在 CONV($\mathrm{R}_{l-1,h}$)中的所有节点。

If $\mathcal{H}_{l-1,h}\neq\emptyset$, then:

- $\mathrm{R}_{l-1,\ h}$ 有一个门生，其定义如下：

$$\mu_{l-1,h}=\arg\min_{j\in\mathcal{H}_{l-1,h}}\left(\sum_{r\in\mathcal{R}_{l-1,h}}\|\mathbf{a}_r-\mathbf{a}_j\|\right),\ h=1,\cdots,m_{l-1}. \tag{11}$$

- In-neighbors of $\mu_{l-1,h}\in\mathcal{L}_l$ is defined by $\mathcal{N}(\mu_{l-1,h})=\mathcal{R}_{l-1,h}$.

注意，如果 $\mathrm{H}_{l-1,\ h}=\emptyset$)，则 $\mathrm{R}_{l-1,\ h}$的门生（记为 $\mu_{l-1,\ h}$）不存在。然后，对于每个 $l\in\mathrm{M}\setminus\{0\}$，$\mathrm{V}_l$聚合 L_{l-1} 所有非空单纯形的门生，定义如下：

$$\mathcal{V}_l=\left\{i\in\mathcal{H}_{l-1,h}:\mathcal{H}_{l-1,h}\neq\emptyset,\ i=\underset{j\in\mathcal{H}_{l-1,h}\neq\emptyset}{\operatorname{argmin}}\left(\sum_{r\in\mathcal{R}_{l-1,h}}\|\mathbf{a}_r-\mathbf{a}_j\|\right),\right.$$
$$\left.h=1,\cdots,m_l-1\right\}. \tag{12}$$

因此，对于每个 $l\in\mathrm{M}\setminus\{0\}$，$\mathrm{V}_l$中的智能体数量满足 $N_l=|\mathrm{V}_l|\leq m_{l-1}$。此不等式成立，是因为 L_{l-1} 中的单纯形并非都必然分配有门生智能体。已知 V_l 和 L_{l-1}后，L_l 由式 (5) 定义。

C. 步骤 *3*：单元分解更新

如果 $\mathrm{H}_{l-1,\ h}\neq\emptyset$，那么，$\mu_{l-1,\ h}$ 存在，且 $(\mathrm{R}_{l-1,\ h})$ 可分解为 $n+1$ 个共享$\mu_{l-1,\ h}$的新单纯形单元。因此，已知 V_l后，前导多胞体可确定性地分解为 m_l 个不同的单纯形，其中$m_l\leq(n+1)m_{l-1}$。

为更清楚地说明，图2展示了如何实现算法1，以根据智能体团队参考配置指定智能体间通信。如图2（a）所示，$CONV(\mathrm{R0},_h)$ 是一个三角形单元，其中至少包含一个智能体，适用于 $\mathrm{h}=1,\cdots,4$。因此，$m_1=12$和 $CONV(\mathrm{R0},_h)$被分解为图2（a）中所示的三个三角形单元，每个单元以蓝色显示。对于第 $l=1$层，$\mathrm{V}_1=\{11,10,6,5\}$，其中11、10、6和5分别是$\mathrm{R}_{0,1}$，$\mathrm{R}_{0,2}$，$\mathrm{R}_{0,3}$，和$\mathrm{R}_{0,4}$，的导师。此外，$m_1=12$；$\mathrm{R1},_1=\{1,2,11\}$、$\mathrm{R1},_2=\{2,12,11\}$、$\mathrm{R1},_3=\{12,1,11\}$、$\mathrm{R1},_4=\{2,3,10\}$、$\mathrm{R1},_5=\{3,12,10\}$、$\mathrm{R1},_6=\{12,2,10\}$、$\mathrm{R1},_7=\{3,4,6\}$、$\mathrm{R1},_8=\{4,12,6\}$、$\mathrm{R1},_9=\{12,3,6\}$、$\mathrm{R1},_{10}=\{4,1,5\}$、$\mathrm{R1},_{11}=\{1,12,5\}$和$\mathrm{R1},_{12}=\{4,5,12\}$定义了图2（b）所示的12个三角形单元的顶点。如图2（b）所示，$CONV(\mathrm{R1},_h)$ 包含一个智能体

算法1 基于参考构造的DNN结构。

获取：智能体的初始位置$\mathbf{a}_1$ 至$\mathbf{a}_N$
获取： 边集 E、 $M = |\mathrm{M}|$以及 V_0 至 V_M。
指定边界智能体 $\mathrm{V}_B = \{b_1, \cdots, b_{N_B}\}$。
根据式（8）指定核心智能体 b_{N_B+1} 。
定义 $\mathrm{V}_0 = \mathrm{V}_B \cup \{b_{N_B+1}\}$。
定义 $\mathrm{L}_0 = \mathrm{V}_0$ 和 $\mathrm{L}_0 = \mathrm{V} \setminus \mathrm{L}_0$。
将前导多胞体分解为 m_0 个单纯形单元
其顶点由 $\mathrm{R}_{0,1}, \cdots$ 和 R_{0,m_0} 定义。
$l = 1$。
当 $\mathcal{L}_{l-1} \neq \emptyset$ **执行**
  $m_l = 0$、 $N_l = 0$ 和 $\mathrm{V}_l = \emptyset$;
  对于 $<h = 1, \cdots, m_{l-1}>$ **执行**
    如果 $CONV(\mathrm{R}_{l-1,h})$ 至少包含一个智能体**则**
      指定 $\mathrm{R}_{l-1,h}$ 的门生，记为 $\mu_{l-1,h}$;
      定义 邻居 of $\mu_{l-1,h}$: $\mathrm{N}(\mu_{l-1,h}) = \mathcal{R}_{l-1,h}$;
      $N_l \leftarrow N_l + 1$;
      $m_l \leftarrow m_l + n + 1$;
      $\mathrm{V}_l = \mathrm{V}_l \cup \{\mu_{l-1,h}\}$;
      指定 $\mathrm{R}_{l,(n+1)(N_l-1)+1}, \cdots,$ 和 $\mathrm{R}_{l,(n+1)N_l}$
    结束 if
  结束 for
  获取 L_l 和 L_l。
  $l \leftarrow l + 1$。
结束 while
$M = l - 1$。

如果$h = 3$，7， 10，11和 $CONV(\mathrm{R}_{1,h})$不包含智能体，则情况相反。因此，$\mathrm{V}_2 = \{7, 8, 9, 13\}$，且7、8、9和13分别是$\mathrm{R}_{1,3} = \mathrm{N}(7)$、$\mathrm{R}_{1,7} = \mathrm{N}(8)$、$\mathrm{R}_{1,10} = \mathrm{N}(9)$以及$\bar{\mathrm{R}}_{1,11} = \mathrm{N}(13)$的门生，适用于 $l = 2$ （见图2（c））。由于$\mathrm{L} = \emptyset$，算法1的while循环在 $l = M = 2$处停止，因此图2（d）所示的深度神经网络指定了智能体间通信。

IV. 训练深度神经网络权重

深度神经网络以完全去中心化且以智能体为中心的方式进行训练，其中每个智能体 $i \in \mathrm{V} \setminus \mathrm{V}0$ 独立优化其局部通信策略。具体而言，智能体 $i \in \mathrm{V} \setminus \mathrm{V}0$ 通过求解一个局部马尔可夫决策过程，为其入邻居 $j \in \mathrm{N}(i)$ 分配自适应通信权重$w_{i,j} \in \mathrm{W}_i$，该局部马尔可夫决策过程的组件和运行分别在第IV-A节和第IV-B节中进行了描述。这种形式无需集中式协调即可实现可扩展的通信感知学习，同时通过第V节建立的理论结果保证覆盖收敛。

A. MDP组件

智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 与一个定义为下式的马尔可夫决策过程相关联

$$\mathcal{M}_i(\mathcal{S}_i, \mathcal{D}_i, \mathcal{A}_i, g_i, P_i, C_i, \gamma_i),$$

其中，M_i 的组件详述如下。

状态集合： 智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 的状态空间定义为 S_i ，并划分为

$$\mathcal{S}_i = \mathcal{S}_i^C \dot{\cup} \mathcal{S}_i^U, \tag{13}$$

其中，S_i^C和 S_i^U分别表示包含 子空间和非包含 子空间。对于每个智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0,$，包含子空间 S_i^C通过离散化由其入邻居 $\mathrm{N}(i)$ 的瞬时位置构成的通信三角形 $\mathrm{T}_i[t]$ 得到。具体而言，

$$\mathcal{S}_i^C = \{s_1, \ldots, s_{M_i^2}\},$$

其中 $M_i = |\mathrm{W}_i|$（见（2）），得到 $|\mathrm{S}_i^C| = M_i^2$。每个状态 $s \in \mathrm{S}_i^C$对应一个三角形单元，其质心为

$$\mathbf{c}_i(s) = \sum_{j \in \mathcal{N}(i)} w_{i,j}(s)\mathbf{r}_j, \qquad s \in \mathcal{S}_i^C, \tag{14}$$

其中重心权重 $w_{i,j} \in \mathrm{W}_i$ 满足（3）。非包含子空间 S_i^U由单个聚合状态表示，该状态涵盖 $\mathrm{T}_i[t]$之外的所有位置。图3展示了由此得到的离散化。

局部目标集合： 令 $\mathbf{d}^j : \mathrm{D} \to \mathbb{R}^2$ 表示环境目标 $j \in \mathrm{D}$ 的位置。智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 可在局部观测到的目标集合定义为

$$\mathcal{D}_i[t] = \{j \in \mathcal{D} : \mathbf{d}_j[t] \in \mathcal{T}_i[t]\}, \qquad \forall i \in \mathcal{V} \setminus \mathcal{V}_0. \tag{15}$$

动作集合： 动作空间 A_i 对 S_i 上的可行转移进行编码，其定义为映射 $\mathrm{A}_i : \mathrm{S}_i \to \mathrm{S}_i$。如果两个三角形单元共享一条边，则认为它们相邻。因此，S_i^C中的每个状态最多有三个相邻单元，因而最多对应四个动作（包括自转移）。若 $s \in \mathrm{S}_i^C$，则 $\mathrm{A}_i(s) \subseteq \mathrm{S}_i^C$；而从 $s \in \mathrm{S}_i^U$出发的动作则转移到包含子空间。图3(b)–(d)展示了这一构造。

目标状态： 选择目标状态 $g_i \in \mathrm{S}_i^C$，使局部目标集合 D_i的覆盖质量最大化。定义

$$\mathbf{h}_i[t] = \begin{cases} \dfrac{1}{|\mathcal{D}_i[t]|} \displaystyle\sum_{j \in \mathcal{D}_i[t]} \mathbf{d}_j[t], & \mathcal{D}_i[t] \neq \emptyset, \\ \dfrac{1}{3} \displaystyle\sum_{j \in \mathcal{N}(i)} \mathbf{r}_j[t], & \mathcal{D}_i[t] = \emptyset, \end{cases} \tag{16}$$

该点在有局部感知目标时表示这些目标的质心，否则表示通信三角形的质心。随后，将目标状态 g_i 定义为包含 $\mathbf{h}_i[t]$的唯一三角形单元。

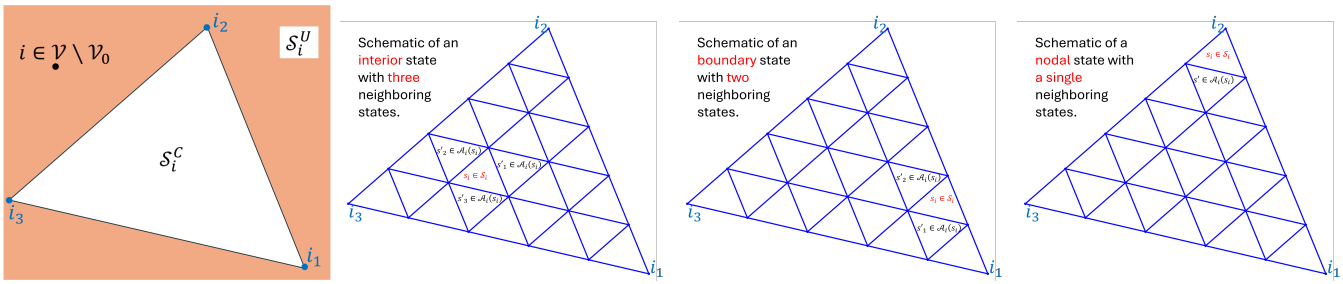
转移动态： 转移核建模为固定基础转移测度的线性组合。令$\Phi_i : \mathrm{S}_i \times \mathrm{A}_i \to \mathbb{R}^d$表示特征映射，并令 $\{\mu_{i,j}\}_{j=1}^d$为图约束基础转移分布。由此得到的转移核为

$$P_i(\cdot \mid s, a) = \sum_{j=1}^{d} \phi_{i,j}(s,a)\,\mu_{i,j}(\cdot), \qquad \forall i \in \mathcal{V} \setminus \mathcal{V}_0. \tag{17}$$

代价函数： 令 $\bar{\mathbf{r}}(s)$ 表示与状态 $s \in \mathrm{S}_i$相关联的三角形单元的质心。该马尔可夫决策过程采用如下状态相关代价函数

$$C_i(s) = \begin{cases} \alpha\left\|\bar{\mathbf{r}}(s) - \bar{\mathbf{r}}(g_i)\right\|, & s \neq g_i, \\ \alpha\left\|\bar{\mathbf{r}}(s) - \bar{\mathbf{r}}(g_i)\right\| - \beta, & s = g_i, \end{cases} \tag{18}$$

其中， $\alpha > 0$ 和 $\beta > 0$ 是设计参数。与 α 成正比的项用于惩罚偏离目标状态 g_i的情况，而终端奖励 β 则鼓励到达目标。



) 图3：(a) $\mathcal{S}_i^U$ 的示意图（以浅红色表示），定义了一个位于智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 的通信三角形之外的单一状态。(b) 内部状态、(c) 边界状态和 (d) 节点状态 $s \in \mathrm{S}_i^C$ 的示意图，它们分别具有三个、两个和一个邻接状态。

折扣因子： 参数 $\gamma_i \in (0, 1)$表示折扣因子。

B. 运行

在时刻 t，智能体 $i \in \mathrm{V} \setminus \mathrm{V}_0$ 可能位于其通信三角形 $\mathrm{T}_i[t]$ 内部或外部。如果智能体位置满足 $\mathbf{r}_i[t] \notin \mathrm{T}_i[t]$，则设定 $M_i = 1$。在此情况下，相应的马尔可夫决策过程状态满足 $s \in \mathrm{S}_i^U$，包含子空间缩减为单元素集合 $|\mathrm{S}_i^C| = 1$，而动作映射 $\mathrm{A}_i(s) = \mathrm{S}_i^C$ 为未包含状态指定唯一的可行后继状态。因此，最优动作是显而易见的，无需进行贝尔曼递推。

相反，当 $s \in \mathrm{S}_i^C$ 时，可行动作集满足 $\mathrm{A}_i(s) \subseteq \mathrm{S}_i^C$。在这种情况下，最优价值函数通过贝尔曼最优性方程计算

$$V_i^*(s) = \min_{a \in \mathcal{A}_i(s)} \left\{ C_i(s,a) + \gamma \sum_{s' \in \mathcal{S}_i} P_i(s' \mid s,a) V_i^*(s') \right\}, \quad s \in \mathcal{S}_i^C, \tag{19}$$

相应的最优策略为

$$\pi_i^*(s) = \arg\min_{a \in \mathcal{A}_i(s)} \left\{ C_i(s,a) + \gamma \sum_{s' \in \mathcal{S}_i} P_i(s' \mid s,a) V_i^*(s') \right\}, \quad s \in \mathcal{S}_i^C. \tag{20}$$

V. 稳定性与收敛

本节将证明所提出的分布式覆盖方法的稳定性和收敛性。

定义 2. 设 $\mathcal{V} = \{b_1, \ldots, b_N\}$。定义

$$\mathbf{y} = \mathrm{vec}\left(\begin{bmatrix} \mathbf{r}_{b_1} & \cdots & \mathbf{r}_{b_N} \end{bmatrix}^T \right) \in \mathbb{R}^{2N}. \tag{21}$$

定义 3。 设 V_l 如式 (4) 所定义。定义

$$\mathbf{y}_l = \mathrm{vec}\left(\begin{bmatrix} \mathbf{r}_{b_{P_{l-1}+1}} & \cdots & \mathbf{r}_{b_{P_l}} \end{bmatrix}^T \right) \in \mathbb{R}^{2N_l}. \tag{22}$$

假设 2：对于任意 $s \in \mathrm{S}_i^U$，目标集 P_i 是一个严格包含于通信三角形 $\mathrm{T}_i(t)$ 内，且与 $\mathrm{T}_i(t)$边对齐的三角形。

假设 3：对于任意 $s \in \mathrm{S}_i^C$，目标集 P_i 是一个严格包含于由最优下一状态 $\pi_i^*(s)$ 诱导的目标三角形内，且与其边对齐的三角形。

定理 1. 假设每个智能体 $b_i \in \mathrm{V} \setminus \mathrm{V}_0$ 均满足 AOC 性质，并且在假设1中，将目标集 P_{b_i} 替换为时变通信三角形 $\mathrm{T}_{b_i}[t]$。则覆盖演化满足

$$\mathbf{y}[t+1] = \mathbf{\Gamma}[t]\mathbf{y}[t], \tag{23}$$

其中，对于所有 t，$\mathbf{\Gamma}[t]$ 均为行随机矩阵。因此，该 (23) 过程定义了一个非齐次马尔可夫过程。

证明。对于每个锚定智能体 $b_i \in \mathrm{V}_0$, $\mathbf{r}_{b_i}[t] = \mathbf{p}_{b_i}$，对所有$t$均成立。对于任意 $b_i \in \mathrm{V} \setminus \mathrm{V}_0$,，AOC 性质和第IV节-B中的转移规则可推出

$$\mathbf{r}_{b_i}[t+1] = \sum_{b_j \in \mathcal{N}(b_i)} \alpha_{b_i,b_j}[t] \mathbf{r}_{b_j}[t],$$
$$\alpha_{b_i,b_j}[t] \geq 0,$$
$$\sum_{b_j \in \mathcal{N}(b_i)} \alpha_{b_i,b_j}[t] = 1.$$

将所有智能体位置堆叠起来，得到式（23），其各项为

$$\Gamma_{i,j}[t] = \begin{cases} 1, & b_i \in \mathcal{V}_0,\ i = j, \\ \alpha_{b_i,b_j}[t], & b_i \in \mathcal{V} \setminus \mathcal{V}_0,\ b_j \in \mathcal{N}(b_i), \\ 0, & \text{otherwise}. \end{cases}$$

$\mathbf{\Gamma}[t]$ 的每一行均非负且元素之和为一；因此，$\mathbf{\Gamma}[t]$是行随机的。□

定理 2。 考虑 $\mathbf{y}[t+1] = \mathbf{\Gamma}[t]\mathbf{y}[t]$，其中每个 $\mathbf{\Gamma}[t] \in \mathbb{R}^{N \times N}$ 均为行随机矩阵。令 V_0 表示满足 $\mathbf{r}_{b_i}[t] \equiv \mathbf{p}_{b_i}$ 的锚定智能体，对所有 $b_i \in \mathrm{V}_0$ 和所有 t均成立。对智能体重新排序后，写作

$$\mathbf{\Gamma}[t] = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B}[t] & \mathbf{A}[t] \end{bmatrix}, \qquad \mathbf{y}[t] = \begin{bmatrix} \mathbf{p} \\ \mathbf{y}_F[t] \end{bmatrix}. \tag{24}$$

假设存在 $T \geq 1$ 和 $\eta \in (0, 1)$，使得：
(C1) 假设 2–3 成立。
(C2) 对每个 t 和每个跟随者索引 i,

$$\sum_{j \in \mathcal{V}_0} \left(\mathbf{\Gamma}[t+T-1] \cdots \mathbf{\Gamma}[t] \right)_{ij} \geq \eta. \tag{25}$$

则跟随者子系统

$$\mathbf{y}_F[t+1] = \mathbf{A}[t]\mathbf{y}_F[t] + \mathbf{B}[t]\mathbf{p} \tag{26}$$

是全局指数稳定的：对所有 $t \geq s$,

$$\|\mathbf{\Phi}(t,s)\|_\infty = \|\mathbf{A}[t-1] \cdots \mathbf{A}[s]\|_\infty \leq (1-\eta)^{\left\lfloor \frac{t-s}{T} \right\rfloor}. \tag{27}$$

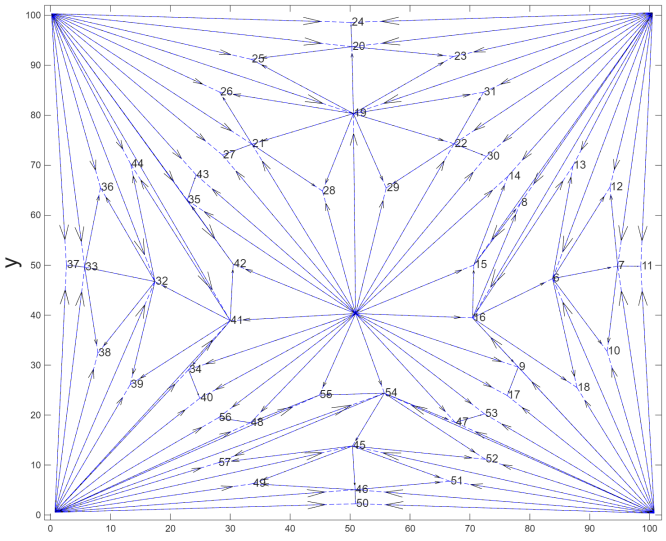


图 4：智能体团队的初始队形和通信链路。

此外，$\mathbf{y}_F[t]$ 收敛，且每个跟随者坐标都收敛到锚点坐标的凸组合。

证明。 由 (24) 和锚点不变性可得，跟随者动力学为 (26)。 $\mathbf{\Gamma}[t]$ 的行随机性意味着$\mathbf{A}[t] \geq 0$ 和$\mathbf{A}[t]\mathbf{1} \leq \mathbf{1}$，即$\mathbf{A}[t]$ 是次随机的。

定义 T步乘积

$$\mathbf{M}[t] = \mathbf{\Gamma}[t+T-1]\cdots\mathbf{\Gamma}[t] = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B}_T[t] & \mathbf{A}_T[t] \end{bmatrix},$$

$$\mathbf{A}_T[t] = \mathbf{A}[t+T-1]\cdots\mathbf{A}[t].$$

由于$\mathbf{M}[t]$ 是行随机的，对于每个跟随者行 $i \in \mathrm{V}\setminus\mathrm{V}_0$,

$$\sum_{j\in\mathcal{V}_0}(\mathbf{M}[t])_{ij} + \sum_{\ell}(\mathbf{A}_T[t])_{i\ell} = 1.$$

根据 (C2)，$\sum_{j\in\mathcal{V}_0}(\mathbf{M}[t])_{ij} \geq \eta$，因此$\sum_{l}(\mathbf{A}_T[t])_{i\ell} \leq 1-\eta$。因此，

$$\|\mathbf{A}_T[t]\|_\infty = \max_i \sum_{\ell}(\mathbf{A}_T[t])_{i\ell} \leq 1-\eta. \tag{28}$$

令 $\mathbf{\Phi}(t,s) = \mathbf{A}[t-1]\cdots\mathbf{A}[s]$。将乘积分组成长度为 T 的块，并利用 $\|\cdot\|_\infty$的次乘性以及 (28)，可得 (27)，从而证明了齐次系统的指数收缩。

展开（26）可得

$$\mathbf{y}_F[t] = \mathbf{\Phi}(t,0)\mathbf{y}_F[0] + \sum_{\tau=0}^{t-1}\mathbf{\Phi}(t,\tau+1)\mathbf{B}[\tau]\mathbf{p}.$$

由于 $[\mathbf{B}[t]\ \mathbf{A}[t]]$ 的各行之和为 1， $\|\mathbf{B}[t]\|_\infty \leq 1$，且（27）意味着 $\|\mathbf{\Phi}(t,\ \tau+1)\|_\infty$ 呈几何衰减，因此该级数收敛；故$\mathbf{y}_F[t]$ 收敛。

最后，每个跟随者的更新都是邻居状态和固定锚点的凸组合，因此每个跟随者坐标始终位于锚点（以及逐渐缩小的跟随者贡献）的凸包内，且其极限是锚点坐标的凸组合。 □

为证明渐近收敛，我们将$\mathbf{p}_i$ 引入为每个智能体 $i \in \mathrm{V}$ 的期望位置；当目标集 D 静止时，该位置为常数。所有锚定智能体 $i \in \mathrm{V}_0$都知道期望位置$\mathbf{p}_i$。相比之下，跟随者智能体 $i \in \mathrm{V}\setminus\mathrm{V}_0$ 无法在本地获取其期望位置。尽管如此，量$\mathbf{p}_i$ 仅作为分析工具，用于刻画跟随者动力学的去中心化收敛。

定义 4 （期望通信三角形）。 对于任意智能体$b_i \in \mathrm{V}\setminus\mathrm{V}_0$,，令 $\mathrm{N}(b_i) = \{b_{i_1},\ b_{i_2},\ b_{i_3}\}$表示其入邻居集合。智能体 b_i 的期望通信三角形定义为

$$\tilde{\mathcal{T}}_{b_i} \triangleq \mathrm{conv}\{\mathbf{p}_{b_{i_1}}, \mathbf{p}_{b_{i_2}}, \mathbf{p}_{b_{i_3}}\}, \tag{29}$$

即入邻居位置的凸包。

定义 5 （诱导目标子集）。 给定环境目标集 D，由 T_{b_i} 覆盖的目标子集定义为

$$\tilde{\mathcal{D}}_{b_i} \triangleq \{j \in \mathcal{D} :\ \mathbf{d}_j \in \tilde{\mathcal{T}}_{b_i}\}. \tag{30}$$

每个智能体 $b_i \in \mathrm{V}\setminus\mathrm{V}_0$ 的期望位置由下式得到

$$\tilde{\mathbf{h}}_{b_i} = \begin{cases} \dfrac{1}{|\tilde{\mathcal{D}}_{b_i}|}\displaystyle\sum_{j\in\tilde{\mathcal{D}}_{b_i}}\mathbf{d}_j[t], & \tilde{\mathcal{D}}_{b_i} \neq \emptyset, \\ \dfrac{1}{3}\displaystyle\sum_{b_j\in\mathcal{N}(b_i)}\mathbf{p}_{b_j}, & \tilde{\mathcal{D}}_{b_i} = \emptyset. \end{cases} \tag{31}$$

算法2通过为智能体集合分配 N 期望位置$\mathbf{p}^{b_1}$、……、$\mathbf{p}^{b_N}$，为环境目标集D提供抽象表示。

定义6。令 $\mathcal{V} = \{b_1,\ \ldots\ldots,\ b_N\}$。定义

$$\mathbf{z} = \mathrm{vec}\left(\begin{bmatrix}\mathbf{p}_{b_1} & \cdots & \mathbf{p}_{b_N}\end{bmatrix}^T\right) \in \mathbb{R}^{2N}. \tag{32}$$

定义7。 令V_l 如式(4)所定义。定义

$$\mathbf{z}_l = \mathrm{vec}\left(\begin{bmatrix}\mathbf{p}_{b_{P_{l-1}+1}} & \cdots & \mathbf{p}_{b_{P_l}}\end{bmatrix}^T\right) \in \mathbb{R}^{2N_l}. \tag{33}$$

定义8。 对于每个智能体 $i \in \mathrm{V}\setminus\mathrm{V}_0$,，令$\mathrm{N}(i)$表示其通信入邻居集合，并令$\mathrm{T}_i$ 表示相应的期望通信三角形。与$\mathbf{h}_i$所包围的单元对应的目标状态 $\tilde{g}_i$中心定义了$\mathbf{p}_i$，并表示为以下凸组合

$$\mathbf{p}_i = \sum_{j\in\mathcal{N}(i)}\tilde{w}_{i,j}\mathbf{p}_j, \qquad \forall i \in \mathcal{V}\setminus\mathcal{V}_0, \tag{34}$$

其中$\tilde{w}_{i,\ j} \in \mathrm{W}_i$。

定义9。 对于每个智能体 $b_i \in \mathrm{V}$，令$\mathrm{N}(b_i) = \{b_{i_1},\ b_{i_2},\ b_{i_3}\}$表示其通信入邻居集合，并令

$$\mathcal{T}_{b_i}[t] = \mathrm{conv}\{\mathbf{r}_{b_{i_1}}[t], \mathbf{r}_{b_{i_2}}[t], \mathbf{r}_{b_{i_3}}[t]\}$$

表示通信三角形。目标状态$g_{b_i} \in \mathrm{S}_{b_i}$的中心对应于一个包围$\mathbf{h}_{b_i}$的单元，记为$\hat{\mathbf{c}}_i$，并表示为凸组合

$$\mathbf{r}_{b_i} = \sum_{j\in\mathcal{N}(b_i)}\hat{w}_{b_i,j}\mathbf{r}_j \tag{35}$$

其中 $\hat{w}_{b_i,\ j} \in \mathrm{W}_i$

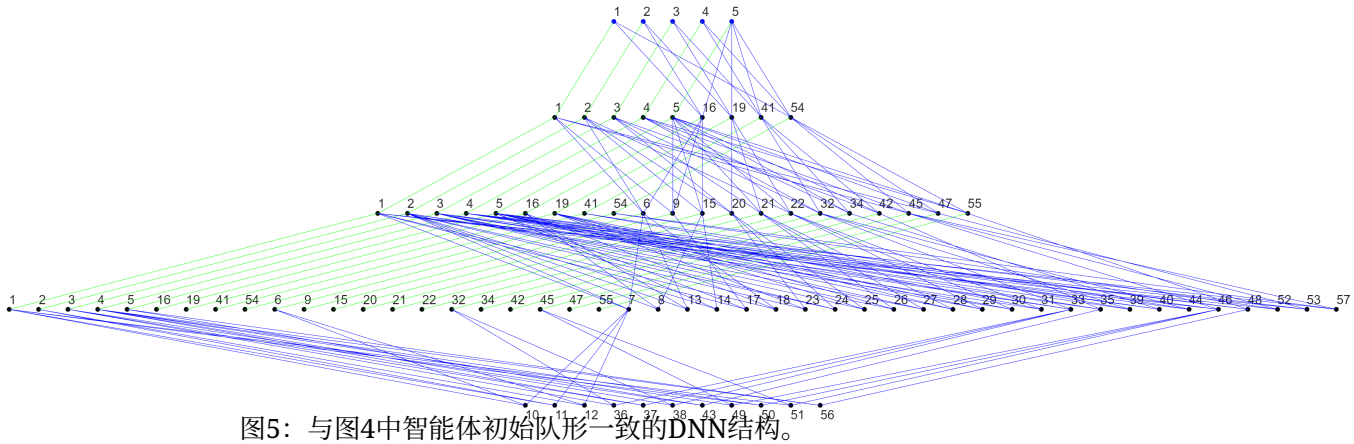


图5：与图4中智能体初始队形一致的DNN结构。

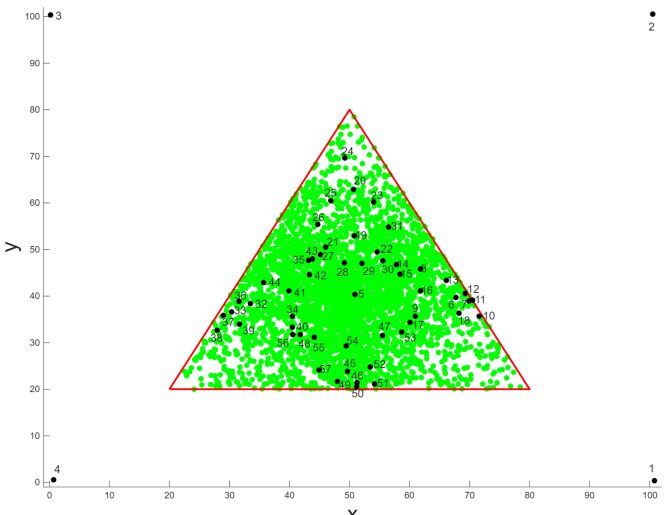


图6：智能体团队的初始队形及通信链路。

定义 10. 我们将 $\tilde{\mathbf{\Gamma}} = [\tilde{\Gamma}_{i,j}]$ 定义为其 $(i,\ j)$ 项

$$\tilde{\Gamma}_{i,j}[t] = \begin{cases} 1, & b_i \in \mathcal{V}_0,\ i = j, \\ \tilde{w}_{b_i,b_j}[t], & b_i \in \mathcal{V} \setminus \mathcal{V}_0,\ b_j \in \mathcal{N}(b_i), \\ 0, & \text{otherwise}. \end{cases} \tag{36}$$

定义11。 我们定义 $\hat{\mathbf{\Gamma}} = [\hat{\Gamma}_{i,\ j}]$ 的$(i,\ j)$项

$$\hat{\Gamma}_{i,j}[t] = \begin{cases} 1, & b_i \in \mathcal{V}_0,\ i = j, \\ \hat{w}_{b_i,b_j}[t], & b_i \in \mathcal{V} \setminus \mathcal{V}_0,\ b_j \in \mathcal{N}(b_i), \\ 0, & \text{otherwise}. \end{cases} \tag{37}$$

$\hat{}$~矩阵 $\mathbf{\Gamma}$、 $\mathbf{\Gamma}$和 $\mathbf{\Gamma}$ 具有相同的严格下块三角结构，在(0，0)位置有一个单位块，且末块列为零。具体而言，对于 $l,\ h \in \mathcal{M}$,

$$\Gamma_{l,h} = \begin{cases} \mathbf{I}, & l = h = 0, \\ \Gamma_{l,h}, & 0 \le h < l \le M, \\ \mathbf{0}, & \text{otherwise}, \end{cases} \tag{38}$$

其中 $\Gamma \in \{\mathbf{\Gamma},\ \hat{\mathbf{\Gamma}},\ \tilde{\mathbf{\Gamma}}\}$。

算法2 用 N 个点表示环境目标

1：获取： 目标集D和$\mathcal{V}_0$的智能体参考位置，记为 $\mathbf{a}^{b_1}$ 至 $\mathbf{a}^{N_0}$，即 DNN结构。
2：获取：智能体的期望位置 $\mathbf{p}^{b_1}$ 至 $\mathbf{p}^{b_N}$。
3: **对于** $<l = 0, \cdots, M>$ **执行**
4: **如果** $l = 0$ **则**
5: **对于** $<i = 1, \cdots, N_0>$ **执行**
6: $\mathbf{p}_{b_i} = \mathbf{a}^{b_i}$。
7: **结束循环**
8: **else**
9: **对于** $<i = P_{l-1}, \cdots, P_l>$ **执行**
10: 赋值 $\tilde{\Gamma}^{b_i}$，根据式 (29)，并 $\tilde{D}^{b_i}$，根据式 (30)。
11: 赋值 $\mathbf{p}_{b_i}$ 根据式 (31)。
12: **结束 for**
13: **结束 if**
14: **结束 for**

命题 1： 给定$\mathbf{z}_0$,智能体团队的期望配置满足

$$\mathbf{z}_l = \sum_{h=0}^{l-1} \tilde{\mathbf{\Gamma}}_{l,h}\mathbf{z}_h, \qquad \forall l \in \mathcal{M} \setminus \{0\}. \tag{39}$$

证明。 根据算法2，每个智能体位置满足

$$\mathbf{p}_{b_i} = \sum_{j \in \mathcal{N}(b_i)} \tilde{w}_{b_i,j}\mathbf{p}_j. \tag{40}$$

将各智能体位置叠加可得

$$\mathbf{z} = \tilde{\mathbf{\Gamma}}\mathbf{z}, \tag{41}$$

由此可直接得到递归关系（39）。 □

定理3。 设每个 AOC智能体 $b_i \in \mathcal{V} \setminus \mathcal{V}_0$ 均能够从与状态 $s \in \mathcal{S}^{b_i}$ 相关联的任意三角形移动到与其最优下一状态 $\pi^*_{b_i}(s) \in \mathcal{S}^{b_i}$相关联的三角形的质心。则对于每个 $b_i \in \mathcal{V} \setminus \mathcal{V}_0$,，期望的实际位置$\mathbf{r}_{b_i}[t]$ 均渐近收敛于$\mathbf{p}^{b_i}$。

证明。 根据定理的假设，子群动力学满足

$$\mathbf{y}_l[t+1] = \sum_{h=0}^{l-1} \mathbf{\Gamma}_{l,h}\mathbf{y}_h[t], \qquad \forall l \in \mathcal{M} \setminus \{0\}. \tag{42}$$

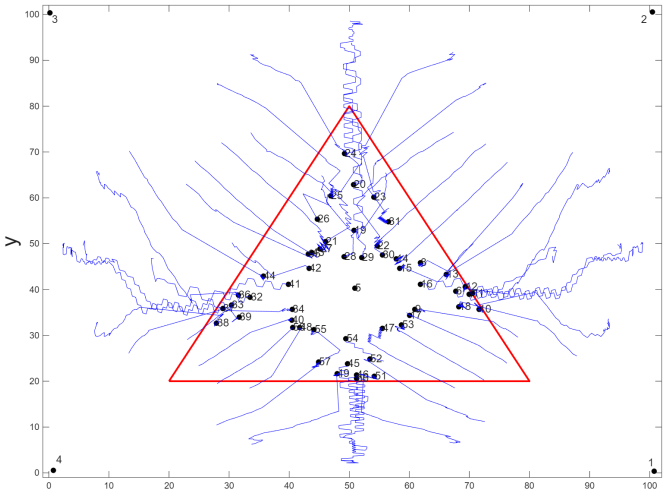


图7：单步可达性假设下的智能体路径。所有智能体 $i \in V$ 都渐近收敛至各自的期望位置 $\mathbf{p}_i$。

对于 $l = 1$，$y_0 = z_0$ 为常数。因此，$D_{b_i} = D_{b_i}$ 和 $\tilde{g}_{b_i} = g_{b_i} \in S(b_i)$ 为所有 $b_i \in V_1$ 定义了固定目标状态。根据第IV节中的马尔可夫决策过程框架，每个智能体 $b_i \in V_1$ 收敛到 p_{b_i}，即 g_{b_i} 所包围的 $h_{b_i} = h_{b_i}$ 的中心，这意味着 $y_1 \to z_1$。假设对于某个 $l \geq 2$，$y_{l-1} = z_{l-1}$。那么，对于所有 $b_i \in V_l$，数据集 $D_{b_i}[t] \to D_{b_i}$ 和 $h_{b_i}[t] \to h_{b_i}$，这意味着相关目标状态 $g_{b_i} \to \tilde{g}_{b_i}$ 收敛。因而，Γ_l，$_h[t] \to \Gamma_l$，$_h$ 对于 $h = 0$，……，$l-1$，且由 (42) 得出 $y_l[t] \to z_l$。根据归纳法，对于所有 $l \in M \setminus \{0\}$，$y_l[t] \to z_l$ 均成立，因此对于所有 $i \in V \setminus V_0$，$r_i[t] \to p_i$。□

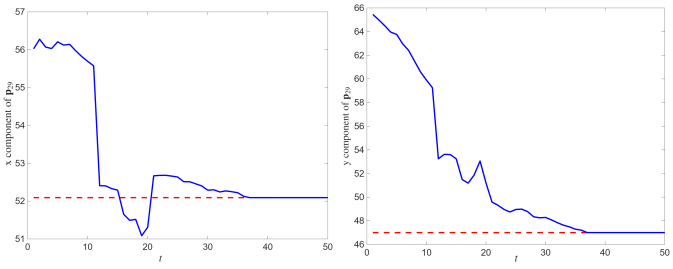
VI. 仿真结果

我们考虑一个57智能体系统，其初始构型如图4所示。根据参考构造，$V_B = \{1,2,3,4\}$ 和 $b_5 = 5$ 分别表示边界智能体和核心智能体，其余所有智能体则归类为内部智能体。由第III节所述框架构建的深度神经网络对所得的智能体间通信结构进行编码，如图5所示（图4中的箭头）。我们希望多智能体系统覆盖图6所示的三角形区域，其中由 D 定义的环境目标数据以绿色显示。智能体团队的期望位置记为 $\mathbf{p}_1$ 至 $\mathbf{p}_{57}$，，以黑色标示。为定义状态空间，我们选择 $M_i = 35$，对每个 $i \in V \setminus V_0$，，这进而意味着 $|S_i^C| = 35^2 = 1225$。

A. 有限时间可达性下的 $\mathbf{h}_i$ 演化

本节假设每个智能体 $i \in V \setminus V_0$ 都能在一个时间步内到达 $\mathbf{h}_i[t]$，这意味着 $P_i = \mathbf{h}_i[t]$。在此假设下，得到的智能体路径如图7所示，其中所有智能体 $i \in V$ 都渐近收敛到其期望位置 $\mathbf{p}_i$。

为进一步说明收敛，图8(a)–(b)展示了智能体29的实际位置和期望位置的 x 分量与 y 分量 $\mathbf{r}_{29}[t]$ 和 $\mathbf{p}_{29}$，随离散时间 t 的变化情况。轨迹表明系统能够快速收敛，$\mathbf{r}_{29}[t]$ 在不到40个时间步内达到 $\mathbf{p}_{29}$。



(a) (b)

) 图8：智能体29实际位置和期望位置的 x 分量与 y 分量随时间的演化，展示了其向 $\mathbf{p}_{29}$ 收敛的过程。

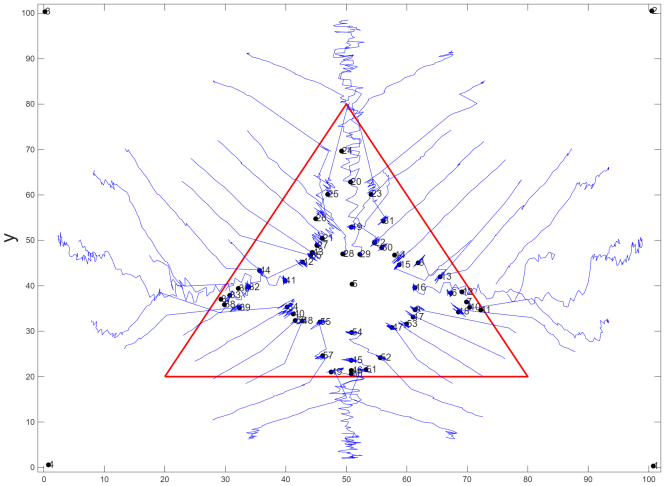


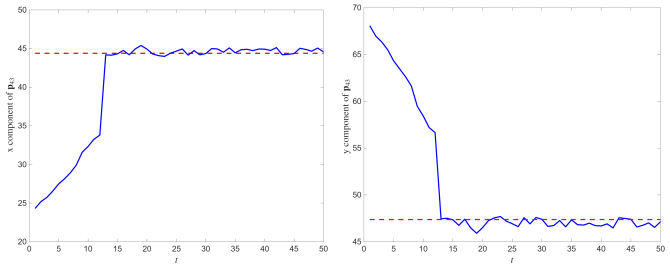
图9：单步可达性假设和假设 3 下的智能体轨迹。所有智能体 $i \in V$ 都渐近收敛到其期望位置 $\mathbf{p}_i$。

B. AOC假设下的演化

本节分析假设 3 下智能体的演化情况。在该假设下，期望位置 $\mathbf{p}_i$ 受限于一个完全包含在 $T_i(t)$ 内的边对齐三角形区域，从而保证对于所有 $i \in V \setminus V_0$，$\eta = 0.05$。在此条件下，图9展示了所有智能体在 x–y 平面上的闭环轨迹，体现了协同运动和空间约束。此外，图10展示了智能体43的期望轨迹的 x 分量与 y 分量随离散时间 t 的变化，说明其能够精确跟踪。

七、结论

本文提出了一种基于结构化学习的框架，用于多智能体系统中的分布式协调和地面覆盖，其中智能体间通信通过几何诱导深度神经网络进行编码。利用参考构造，将智能体系统地划分为边界组和内部组，从而形成一种分层通信架构，其通信权重受到明确约束且具有可解释性。这些权重从有限集合中选取，并由分布式马尔可夫决策过程控制，从而确保交互经过归一化



(a) (b) 图10：智能体43的实际位置和期望位置的 x 分量与 y分量随时间的变化，展示了其收敛至$\mathbf{p}_{43}$的过程。

且局部决策过程定义良好。在此框架下，证明了在 AOC 假设下，智能体轨迹将收敛至与环境目标数据相关联的期望目标配置。数值仿真验证了所提出的基于策略的分散式覆盖策略，并展示了其捕捉几何结构及实现复杂区域有效覆盖的能力。

参考文献

[1] Farshid Abbasi、Afshin Mesbahi 和 Javad Mohammadpour Velni。基于沃罗诺伊图的新型移动传感器网络全覆盖控制方法。*IEEE*控制系统技术汇刊，27(1):409–417，2017。[2] Onur Arslan 和 Daniel E. Koditschek。异构圆盘形机器人的基于沃罗诺伊图的覆盖控制。*IEEE*机器人学汇刊，29(2):395–411，2013。[3] G¨okhan M Atınc.、Du˘san M Stipanovi´c 和 Petros G Voulgaris。基于群体的方法：多智能体系统的动态覆盖控制。*Automatica*，112:108637，2020。[4] Yang Bai、Yujie Wang、Mikhail Svinin、Evgeni Magid 和 Ruisheng Sun。具有避障功能的自适应多智能体覆盖控制。*IEEE*控制系统快报，6:944–949，2021。[5] Shiba Biswal、Karthik Elamvazhuthi 和 Spring Berman。利用局部密度反馈的多智能体系统分散式控制。*IEEE*自动控制汇刊，67(8):3920–3932，2021。[6] William Coffey、Matthew A. McCourt 和 Magnus Egerstedt。利用沃罗诺伊镶嵌实现异构覆盖与多资源分配。载于*IEEE*国际机器人与自动化会议 *(ICRA)*，第11342–11348页，2023。[7] Jorge Cort´es。机器人传感器网络的覆盖优化与空间负载均衡。*IEEE*自动控制汇刊，55(3):749–754，2010。[8] Anna Dai、Rongpeng Li、Zhifeng Zhao 和 Honggang Zhang。用于无人机覆盖控制的图卷积多智能体强化学习。载于*2020*年国际无线通信与信号处理会议 *(WCSP)*，第1106–1111页。IEEE，2020。[9] Ahmad Din、Muhammed Yousoof Ismail、Babar Shah、Mohammad Babar、Farman Ali 和 Siddique Ullah Baig。智慧农业中基于深度强化学习的多智能体区域覆盖控制。计算机与电气工程，101:108089，2022。[10] Alireza Dirafzoon、Mohammad Bagher Menhaj 和 Ahmad Afshar。具有非线性动力学的多智能体系统分散式覆盖控制。*IEICE*信息与系统汇刊，94(1):3–10，2011。[11] Karthik Elamvazhuthi 和 Spring Berman。机器人群体基于扩散的覆盖的非线性推广。载于*2018*年*IEEE*决策与控制会议 *(CDC)*，第1341–1346页。IEEE，2018。[12] YuZe Feng、Gang Lu、WenJie Bai、JianHang Zhao、YuMo Bai 和 Tao Xu。基于 k 均值算法的多智能体系统快速覆盖控制。载于*2020*年第7届国际信息、控制论与计算社会系统会议 *(ICCSS)*，第870–874页，2020。[13] Jia Hua、Yifan Xu 和 Xiaoming Hu。异构机器人团队的持续监视与覆盖控制。*IEEE*机器人学汇刊，2025。提前发表。

[14] Solmaz S. Kia 和 Jorge Cort´es。具有时变密度函数的覆盖控制分布式算法。*Automatica*，89 220，2018 210: – 。[15] Hyeongseok Kim、Marco Santos、Jose Guerrero-Bonilla、Anthony Yezzi 和 Magnus Egerstedt。面向时间敏感型应用的不同最高速度移动机器人的覆盖控制。*IEEE*机器人与自动化快报，7(2):3136–3143，2022。[16] Vishaal Krishnan 和 Sonia Mart´ınez。多智能体覆盖控制算法的多尺度分析。*Automatica*，145:110516，2022。[17] Martin Lauer 和 Martin Riedmiller。合作式多智能体系统分布式强化学习算法。载于国际机器学习会议 *(ICML)*，第535–542页，2000。[18] Wenhao Luo 和 Katia Sycara。具有连通性维护功能的机器人传感器网络基于沃罗诺伊图的覆盖控制。载于*2019*年国际多机器人与多智能体系统研讨会 *(MRS)*，第148–154页。IEEE，2019。[19] Minh Tri Nguyen、Luis Rodrigues、Cristina Stoica Maniu 和 Sorin Olaru。动态多智能体分散式覆盖的离散化最优控制方法。载于*2016*年*IEEE*国际智能控制研讨会 *(ISIC)*，第1–6页。IEEE，2016。[20] Reza Olfati-Saber。嵌入一致性滤波器的分布式卡尔曼滤波器。*IEEE*决策与控制会议论文集，2002。用于分布式感知与覆盖的早期MDP式决策理论表述。[21] Shivang Patel、Senthil Hariharan、Pranav Dhulipala、Ming C Lin、Dinesh Manocha、Huan Xu 和 Michael Otte。城市环境中的多智能体覆盖。*arXiv*预印本 *arXiv:2008.07436*，2020。[22] Arash Rahmani 和 Mehran Mesbahi。基于学习的多智能体系统分散式控制。*IEEE*网络系统控制汇刊，6(3):1027–1038，2019。[23] Hamed Sadeghi、Mahdi Sadraddini 和 Calin Belta。异构机器人对多种事件类型的覆盖控制。载于*IEEE*国际机器人与自动化会议 *(ICRA)*，第7992–7998页，2019。[24] Marco Santos 和 Magnus Egerstedt。具有异构感知能力的多机器人团队覆盖控制。载于*IEEE/RSJ*国际智能机器人与系统会议 *(IROS)*，第5856–5861页，2018。[25] Esmaeil Seraj 和 Matthew Gombolay。无人机协同控制，用于以人为中心的野火主动感知。载于*2020*年美国控制会议 *(ACC)*，第1645–1652页，2020。[26] Cheng Song、Gang Feng、Yuan Fan 和 Yong Wang。多智能体网络的分散式自适应态势感知覆盖控制。*Automatica*，47(12):2749–2756，2011。[27] Jian Xiao、Gang Wang、Ying Zhang 和 Lei Cheng。一种基于强化学习的分布式多智能体动态区域覆盖算法。*IEEE Access*，8:33511–33521，2020。[28] Ying Zhang 和 Sonia Mart´ınez。网络化系统中的分布式学习与决策。*Automatica*，125:109417，2021。[29] Meng Zhong 和 Christos G. Cassandras。移动智能体的分布式覆盖控制与数据收集。*IEEE*自动控制汇刊，56(10):2449–2464，2011。



Hossein Rastgoftar 现任亚利桑那大学助理教授。此前，他于2020年至2021年担任密歇根大学兼职助理教授。他还曾于2017年至2020年担任密歇根大学安娜堡分校航空航天工程系助理研究科学家，并于2015年至2017年担任博士后研究员。他获得了伊朗设拉子设拉子大学机械工程—热流体专业的理学学士学位，在设拉子大学和美国佛罗里达州奥兰多的中佛罗里达大学获得了机械系统和固体力学理学硕士学位，并获得了2015年，他在费城德雷塞尔大学获得机械工程博士学位。