arXiv:2301.09689v1 [cs.MA] 23 Jan 2023

用于去中心化多智能体周界防御的图神经网络

Elijah S. Lee [1,*], **Lifeng Zhou** [2], **Alejandro Ribeiro** [1] **和 Vijay Kumar** [1]

[1]宾夕法尼亚大学 *GRASP* 实验室，美国宾夕法尼亚州费城[2] 德雷塞尔大学电气与计算机工程系，美国宾夕法尼亚州费城

通讯作者*：Elijah S. Lee
elslee@seas.upenn.edu

摘要

本文研究去中心化多智能体周界防御问题，该问题要求根据防御者的局部感知和通信来计算其动作，以最大限度地捕获入侵者。实际应用面临的一项主要挑战，是如何使周界防御策略能够扩展到大规模问题实例。为此，我们利用图神经网络（GNN）开发了一种模仿学习框架，学习从防御者的局部感知及其通信图到其动作的映射。所提出的基于 GNN 的学习网络通过模仿集中式专家算法进行训练，使学习到的动作接近专家算法生成的动作。我们展示了，与其他基线算法相比，所提出的网络表现更接近专家算法，并且能捕获更多入侵者，因而性能更优。我们的基于 GNN 的网络在小规模场景下训练，并可推广至大规模场景。我们在具有不同团队规模和配置的周界防御博弈场景中进行实验，以展示所学习网络的性能。

关键词：图神经网络、周界防御、多智能体系统、感知—行动—通信回路、模仿学习

1 引言

周界防御博弈问题研究这样一种场景：防御者受限于沿周界移动，并试图捕获入侵者；而入侵者则试图在不被防御者捕获的情况下抵达周界（Shishika 和 Kumar，2020年）。此前已有不少研究通过在平面博弈空间中展开对抗来解决这一问题（Shishika 和 Kumar，2018年；Chen 等人，2021年）。然而，在现实世界中，周界可能呈三维形状，因为参与者（例如防御者和入侵者）可能具备进行三维运动的能力。例如，防御者试图保护的建筑物周界可能由一个半球围成。因此，防御机器人应能够在三维空间中移动。例如，空中机器人已在多种场景中得到广泛研究（Chen 等人，2020年；Nguyen 等人，2019年；Lee 等人，2016年、2020a），这些场景都可以成为周界防御的现实应用案例。例如，入侵者试图袭击森林中的军事基地，而防御者则力图捕获入侵者。

在这项工作中，我们研究多个智能体协作完成任务的领域中的周界防御问题。多智能体协作已在许多领域得到探索，包括环境测绘（Liu 等人，2022年；Thrun 等人，2000年）、搜索与救援（Miller 等人，2020年；Baxter 等人，2007年）、

目标跟踪（Ge等人，2022年；Lee 等人，2022b）、按需无线基础设施（Mox等人，2020年）、交通运输（Xu等人，2022年；Ng等人，2022年）以及多智能体学习（Kim等人，2021年）。我们采用一支机器人团队，共同致力于守卫周界这一目标。我们着重为防御者团队开发去中心化策略，原因有以下几点：(i) 队友可以动态加入或退出，而不会破坏明确的层级结构；(ii) 集中式系统可能难以应对团队联合状态空间的高维度；(iii) 防御者的通信范围有限，因此只能进行局部通信。

为此，我们旨在开发一个框架，使一组防御者能够基于局部感知和通信，采用去中心化策略协作防御周界。具体而言，我们探索基于学习的方法，通过模仿专家算法（如最大匹配算法（Chen 等人，2014））来学习策略。最大匹配算法通过穷举搜索寻找最优策略；由于这种方法本质上是组合式的，并且假设拥有全局信息，因此在大规模问题中计算开销极高。我们采用图神经网络作为学习范式，并证明训练后的网络能够达到接近专家算法的表现。图神经网络具有去中心化通信架构，能够捕获邻近交互；其迁移能力使其能够泛化到此前未见过的场景（Ruiz 等人，2021）。我们证明，所提出的基于图神经网络的网络能够泛化到大规模问题，从而解决多机器人周界防御博弈。

基于这一认识，我们在本文中作出以下主要贡献：

基于图神经网络的去中心化周界防御框架。 我们提出了一种新颖的学习框架，利用基于图的表示来建模周界防御博弈。据我们所知，我们是首个通过图神经网络学习去中心化策略，从而解决去中心化半球周界防御问题的研究团队。

兼具可扩展性的稳健周界防御性能。 我们证明，我们的方法表现接近专家策略（即 Chen 等人（2014）的最大匹配算法），并优于其他基线算法。我们提出的网络在小规模场景下训练，并可泛化至大规模场景。

2 相关工作

2.1 周界防御

在周界防御博弈中，防御者沿周界移动以捕获入侵者，而入侵者则试图在不被防御者捕获的情况下抵达周界。详细综述请参见（Shishika 和 Kumar，2020年）。许多先前研究探讨了平面博弈空间中的交战（Shishika 和 Kumar，2018年；Macharet等人，2020年；Chen 等人，2021年；Bajaj等人，2021年；Hsu等人，2022年）。例如，（Shishika 和 Kumar，2018年）在平面博弈空间中解决了合作式多人周界防御博弈。此外，（Macharet等人，2020年）提出了一种基于入侵者到达时间估计的自适应分区策略。随后，（Chen 等人，2021年）提出将周界防御问题建模为流网络中的一个实例。此外，（Bajaj等人，2021年）讨论了圆锥环境中的交战，而（Hsu等人，2022年）则研究了异构团队模型。

近期已有研究探讨周界防御问题的高维扩展（Lee and Bakolas, 2021年；Yan 等人, 2022年；Lee 等人, 2020b, 2021年, 2022a）。例如，Lee and Bakolas（2021年）分析了在高维欧几里得空间中，防御者守卫一个闭凸目标集以抵御攻击者的双人微分博弈。

Yan 等人（2022年）研究了一种三维多人追逃博弈，其中多名追捕者保护目标区域，抵御多名逃避者。Lee 等人（2020b, 2021年, 2022a）考虑了空中防御者与地面入侵者之间进行的游戏。

上述所有工作都着重于解决集中式周界防御问题，这类问题假设参与者能够全局了解其他参与者的状态。然而，当参与者数量增多时，分散式控制就成为必需。为解决这一问题，Velhal 等人（2022年）将周界防御博弈表述为凸形状周界上的一个去中心化多机器人时空多任务分配问题。Paulos 等人（2019年）提出了一种神经网络架构，用于在单位圆周界上训练分散式智能体策略，其中防御者的动作空间简单且为二元。与上述工作不同，我们关注高维周界，具体而言是半球周界，且动作空间为连续空间。我们通过图神经网络学习去中心化策略，以解决多智能体周界防御问题。

2.2 图神经网络

我们采用图神经网络作为学习范式，因为它具有理想的特性：其分散式架构能够捕获邻近智能体之间的交互，而迁移能力则使其能够泛化到此前未见的情形（Gama 等人，2019年；Ruiz 等人，2021年）。此外，图神经网络已在各种多机器人问题中取得显著成效，例如编队控制（Tolstaya 等人，2019年）、路径规划（Li 等人，2021年）、任务分配（Wang 和 Gombolay，2020年）以及多目标跟踪（Zhou 等人，2021年；Sharma 等人，2022年）。特别是，Tolstaya 等人（2019年）通过模仿一个利用全局信息的集中式编队控制器，使用图神经网络为移动机器人群学习分散式集群行为。随后，Li 等人（2021年）采用图神经网络，为多个机器人规划从起始位置到目标位置的无碰撞路径，应用场景为障碍物密集环境。他们证明，仅凭局部观测和邻近通信，其分散式路径规划器便能达到接近专家水平的性能，并且还能泛化到规模更大的机器人网络。基于图神经网络的方法还被用于学习多机器人任务调度场景（Wang 和 Gombolay，2020年）以及多目标跟踪场景（Zhou 等人，2021年；Sharma 等人，2022年）中的组合优化问题的解。

3 问题表述

3.1 研究动机

周界防御是一个相对较新的研究领域，近年来才受到关注。其中一个特殊挑战是，高维周界增加了防御者执行最优策略时所面临的空间和算法复杂性。尽管许多以往研究都考虑了平面博弈空间中的交锋，并推导出二维运动下的最优策略，但在现实场景中实际应用周界防御博弈时，向高维空间扩展不可避免。例如，防御者试图保护的建筑物周界可以由一般形状围成，例如半球。由于防御者无法穿过建筑物，并且假设他们始终靠近建筑物，因此他们需要沿穹顶表面移动，由此形成“半球周界防御博弈”。入侵者在半球的底平面上移动，这意味着其移动过程中高度恒定。在现实世界中，假设入侵者可能希望保持较低的高度以躲避防御者，因此其运动被限制在二维空间内。

值得注意的是，与双方智能体都能在三维空间中自由移动的问题相比，半球防御问题更难求解。此前已有研究允许防御者和入侵者都在三维空间中移动（Yan 等人，2022年、2019年、2020年）。在所有这些情形中，作者都能够明确推导出最优解，即使是在多机器人场景中也是如此。尽管我们的问题将防御者的动力学限制在半球表面上，但这些约束使得寻找最优解变得难以处理且颇具挑战。

3.2 半球周界防御

我们考虑一个半径为 R 的半球形穹顶作为周界。设置半球约束，是为了让防御者能够安全地沿周界（例如建筑物）移动。在这个游戏中，考虑两组参与者：$\mathbf{D} = \{D_i\}_{i=1}^{N}$ 表示 N 个防御者，$\mathbf{A} = \{A_j\}_{j=1}^{N}$ 表示 N 个入侵者。防御者 D_i 受限于在穹顶表面移动，而入侵者 A_j 受限于在地平面上移动。当索引 i 和 j 无关时，我们将省略它们。图 1 展示了一个 10 对 10 周界防御实例。参与者在球坐标中的位置为：$\mathbf{z}_D = [\psi_D, \phi_D, R]$ 和 $\mathbf{z}_A = [\psi_A, 0, r]$，其中 ψ 和 ϕ 分别是方位角和仰角，由此得到相对位置：$\mathbf{z} \triangleq [\psi, \phi, r]$，其中 $\psi \triangleq \psi_A - \psi_D$ 和 $\phi \triangleq \phi_D$。参与者的位置也可以用笛卡尔坐标表示为：$\mathbf{x}_D$ 和 $\mathbf{x}_A$。所有智能体均以单位速度移动；防御者通过缩短距离至 ϵ 以内来捕获入侵者，且防御者和入侵者都会在捕获过程中被消耗。若入侵者在时间 t_f 到达周界（即 $r(t_f) = R$），且未被任何防御者捕获（即$||\mathbf{x}^{A_i}(t) - \mathbf{x}^{D_j}(t)|| > \epsilon, \forall D_j \in \mathbf{D}\ \forall t < t_f$），则入侵者获胜。防御者若捕获入侵者，或无限期地阻止其得分（即 $\phi(t) = \psi(t) = 0, r(t) > R$），则获胜。本研究的主要目标是，在给定一组初始配置的情况下，最大化防御者的捕获次数。

3.3 最优突破点

给定$\mathbf{z}_D$、$\mathbf{z}_A$，我们将入侵者试图到达目标的周界上的点称为突破点，如图2 B 所示。我们将确定突破点的方位角称为突破角，记作 θ；将$(\mathbf{z}_A - \mathbf{z}_B)$与 B 处切线之间的夹角称为接近角，

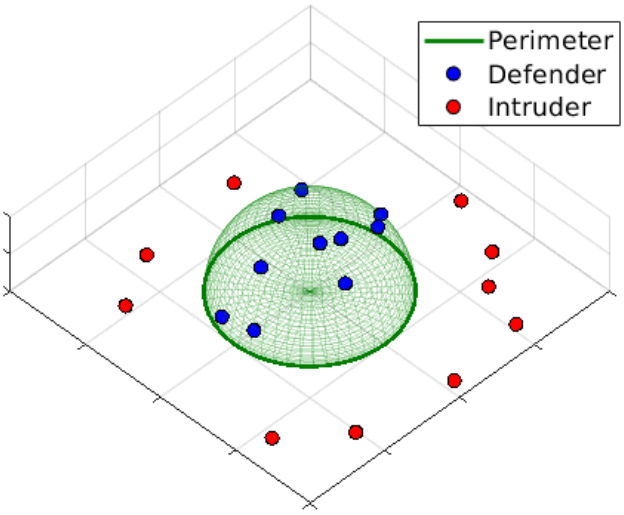


图1. 10对10周界防御实例。防御者只能在穹顶表面移动，而入侵者只能在平面图上移动。

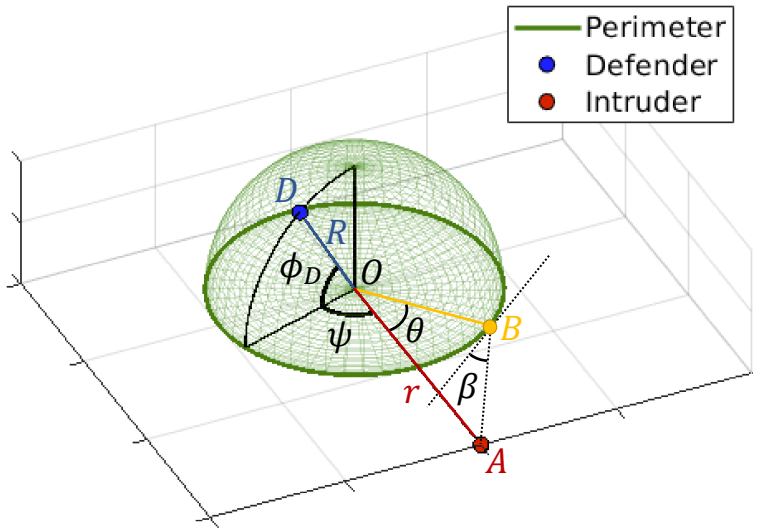


图2. 一对一半球防御博弈中的坐标和相关变量。

记作 β。Lee 等人（2020b）证明，给定防御者$\mathbf{z}_D$ 和入侵者$\mathbf{z}_A$ 作为点粒子的当前位置，存在一个唯一的突破点，使得防御者和入侵者的最优策略都是朝该点移动，这个点称为最优突破点。与最优突破点对应的突破角和接近角分别称为最优突破角，记作θ^*，以及最优接近角，记作 β^*。如 Lee 等人（2020b）所述，尽管 θ^* 和 β^* 不存在闭式解，但可以通过求解两个控制方程随时计算得出：

$$\beta^* = \cos^{-1}\left(\nu \frac{\cos\phi_D \sin\theta^*}{\sqrt{1-\cos^2\phi_D\cos^2\theta^*}}\right) \tag{1}$$

and

$$\theta^* = \psi - \beta^* + \cos^{-1}\left(\frac{\cos\beta^*}{r}\right) \tag{2}$$

3.4 目标到达时间与收益函数

我们将到达 B 所需的时间称为目标到达时间，并将 $\tau_D(\mathbf{z}_D,\mathbf{z}_B)$ 定义为防御者目标到达时间，将 $\tau_A(\mathbf{z}_A,\mathbf{z}_B)$ 定义为入侵者目标到达时间，并将以下内容定义为收益函数：

$$p(\mathbf{z}_D,\mathbf{z}_A,\mathbf{z}_B) = \tau_D(\mathbf{z}_D,\mathbf{z}_B) - \tau_A(\mathbf{z}_A,\mathbf{z}_B) \tag{3}$$

如果 $p<0$，则防御者更快到达 B；如果 $p>0$，则入侵者更快到达 B。因此，防御者旨在最小化 p，而入侵者旨在最大化它。

3.5 最优策略与纳什均衡

（Lee 等人，2020b年）证明，防御者和入侵者的最优策略是在任何时刻都以各自的最大速度朝最优突破点移动。令 Ω 和 Γ 为导向 B 的连续 v_D 和 v_A，使得 τ_D（$\mathbf{z}_D$，Ω）≙ τ_D（$\mathbf{z}_D$，$\mathbf{z}_B$），且 τ_A（$\mathbf{z}_A$，Γ）≙ τ_A（$\mathbf{z}_A$，$\mathbf{z}_B$）；再令 Ω* 和 Γ* 为分别使 τ_D（$\mathbf{z}_D$，Ω）和 τ_A（$\mathbf{z}_A$，Γ）最小化的最优策略，则游戏中的最优性

可表示为纳什均衡：

$$p(\mathbf{z}_D, \mathbf{z}_A, \Omega^*, \Gamma) \leq p(\mathbf{z}_D, \mathbf{z}_A, \Omega^*, \Gamma^*) \leq p(\mathbf{z}_D, \mathbf{z}_A, \Omega, \Gamma^*) \tag{4}$$

3.6 问题定义

为了在 N 对 N 防御中最大化捕获数量，我们首先回顾 1 对 1 周界防御博弈的动力学。Lee 等人（2020b）证明，在一对一博弈中，防御者的最佳行动是朝最优突破点 移动（定义见第 3.3 节）。如果收益 p（定义见第 3.4 节）为负，防御者会比入侵者更快到达最优突破点；如果 $p > 0$，则入侵者会更快到达。由此我们推断，在 N 对 N防御中最大化捕获数量，等同于寻找防御者与入侵者之间的一种匹配，使匹配双方的负收益数量达到最大。在最优匹配中，整个博弈过程中的负收益数量保持不变，因为防御者与入侵者配对后的每场博弈均达到纳什均衡（见第 3.5 节）。

专家分配策略是一种最大匹配（Shishika 和 Kumar，2018；Chen 等人，2014）。为执行该算法，我们将**D** 和**A** 作为两组节点（即 $\mathcal{V} = \{1, 2, .., N\}$），构建一个二分图，并将防御者与入侵者之间可能的分配定义为边。对于**D**中的每个防御者/节点 D_i，我们找出该防御者能够合理感知的**A** 中的所有入侵者/节点 A_j，并计算所有配对对应的收益 p_{ij}。如果 $p_{ij} < 0$，我们就称 D_i 被强分配给 A_j。利用最大匹配得到的边集 $\mathcal{E}$，我们可以使强分配配对的数量最大化。为保证唯一性，我们选择使游戏的值最小的匹配，其定义为

$$V = \sum_{(D_i, A_j) \in \mathcal{E}^*} p_{ij}, \tag{5}$$

其中，$\mathcal{E}^*$是 $\mathcal{E}$中收益为负的子集（即 $\mathcal{E}^* = \{(D_i,\ A_j) \in \mathcal{E} \mid p_{ij} < 0\}$）。这种唯一分配可确保尽可能早地最大化捕获数量。然而，随着团队规模增大，使用最大匹配算法进行穷举搜索的代价可能非常高。该方法本质上是组合式的，并假设具备集中式信息和完全可观测性。相反，我们旨在寻找利用局部感知 $\{\mathcal{Z}_i\}_{i \in \mathcal{V}}$ 的去中心化策略（见第 4.1 节）。为此，我们将本文的主要问题形式化如下。

问题 1（基于图神经网络的去中心化周界防御）。设计一个基于图神经网络的学习框架，学习从防御者的局部感知 $\{\mathcal{Z}_i\}_{i \in \mathcal{V}}$ 及其通信图 $\mathcal{G}$到其动作 $\mathcal{U}$的映射 $\mathcal{M}$，即 $\mathcal{U} = \mathcal{M}\left(\{\mathcal{Z}_i\}_{i \in \mathcal{V}}, \mathcal{G}\right)$，使其尽可能接近由集中式专家算法选定的动作集 $\mathcal{U}^g$。

我们描　述我们用于解决问题 1 的学习架构，具体内容将在下一节介绍。

4 方法

本文使用图神经网络学习用于周界防御的去中心化策略。我们的方法可实时推理，并能够扩展至大量智能体。我们采用专家分配策略来训练一支通过通信通道共享信息的防御者团队。第 4.1 节介绍用于处理输入图神经网络特征的感知模块。通过图神经网络学习去中心化算法，以及为防御者规划候选匹配，将在

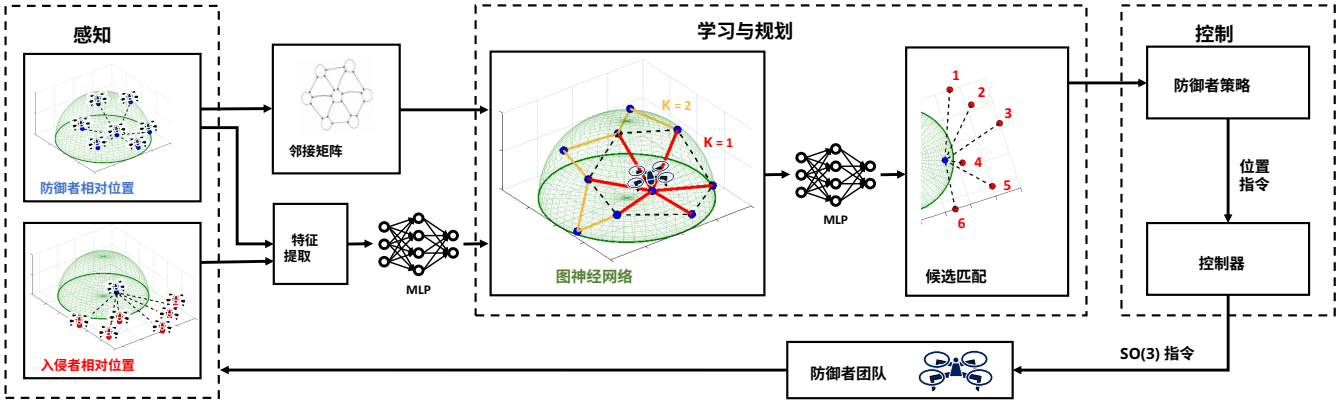


图 3. 总体框架。感知模块收集局部信息。学习与规划模块通过 K 跳邻近通信，使用图神经网络处理收集到的信息。控制模块计算最优策略，并执行控制器以形成闭环。

第 4.2 节讨论。第 4.3 节说明防御者团队的控制，第 4.4 节详述训练流程。整体框架如图 3 所示。在架构选择上，我们将控制模块与学习框架解耦，因为直接学习动作并无必要。只需学习智能体间的分配，便可通过最优策略计算出最佳动作（第 3.5 节）。

4.1 感知

在本节中，我们假设有 N 个空中防御者和 N 个地面入侵者。每个防御者 D_i 都配备了一个传感器，面朝周界外侧，视场为 FOV。假设使用鱼眼相机，防御者的水平视场 FOV 设为 π。

4.1.1 入侵者特征

对于每个 i，防御者会观测到入侵者集合 A_j，而 D_i 与 A_j 之间的球坐标相对位置表示为 $\mathcal{Z}_i^A = \{\mathbf{z}_{ij}^A\}_{j \in N_A^f}$，其中 N_A^f 是入侵者特征的数量。输入特征的数量 N_A^f 和 N_D^f 分别选为最近检测到的入侵者和邻近智能体的固定数量。尽管防御者能够检测到感知范围内任意数量的入侵者，但系统会选取固定数量的检测结果，以确保系统可扩展。在分散式环境中，防御者应能根据其局部感知决定采取何种行动。我们通过实验将固定数量设为 10，因为专家算法（即最大匹配）总会将一名防御者分配给距离最近的 10 个入侵者中的一个机器人。

4.1.2 防御者特征

为了使系统具有可扩展性，我们构建了与固定数量最近防御者的通信。每个防御者 D_i 都会与其通信范围 r_c 内的附近防御者 D_j 通信。对于每个 i，D_i 与 D_j 之间的相对位置表示为 $\mathcal{Z}_i^D = \{\mathbf{z}_{ij}^D\}_{j \in N_D^f}$，其中 N_D^f 是防御者特征的数量。选择 3 是因为，与许多其他机器人通信会使每个防御者都能获得环境的完整信息（即集中式信息），而在假设机器人分散分布的情况下，3 是机器人能够收集各个方向信息的最小数量。如果检测到的入侵者少于 10 个，或邻近防御者少于 3 个，我们会提供虚拟值来填充感知

输入矩阵。保持输入特征数量恒定很重要，因为神经网络无法处理大小不一的特征。

4.1.3 特征提取

通过拼接观测到的入侵者和已通信防御者的相对位置来执行特征提取，形成局部感知 $\mathcal{Z}_i = \{\mathcal{Z}_i^A, \mathcal{Z}_i^D\}$。提取出的特征被输入多层感知机（MLP），以生成后处理特征向量 $\mathbf{x}_i$，该向量将通过通信在邻居之间交换。

4.2 学习与规划

我们采用通过 K 跳通信的图神经网络。防御者将其感知到的特征与邻近机器人进行通信。通信图 $\mathcal{G}$ 通过连接通信范围 r_c 内彼此接近的防御者而形成，所得邻接矩阵 $\mathbf{S}$ 被输入图神经网络。

4.2.1 图移位操作

我们考虑每个防御者 i，$i \in \mathcal{V}$都有一个特征向量 $\mathbf{x}_i \in \mathbb{R}^F$，用于表示来自 D_i 的后处理信息。收集所有防御者的特征向量 $\mathbf{x}_i$ 后，我们得到防御者团队的特征矩阵 $\mathbf{D}$：

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^\mathsf{T} \\ \vdots \\ \mathbf{x}_N^\mathsf{T} \end{bmatrix} = [\mathbf{x}^1, \cdots, \mathbf{x}^F] \in \mathbb{R}^{N \times F}, \tag{6}$$

其中，$\mathbf{x}^f \in \mathbb{R}^N$，$f \in [1, \cdots, F]$ 是所有防御者的特征 f的集合；即，$\mathbf{x}^f = [\mathbf{x}_1^f, \cdots, \mathbf{x}_N^f]^\mathsf{T}$，其中 $\mathbf{x}_i^f$表示 D_i，$i \in \mathcal{V}$的特征 f。我们对每个 D_i执行图移位操作，将其邻居的特征线性组合，即 $\sum_{j \in \mathcal{N}_i} \mathbf{x}j$。因此，对于图$\mathcal{G}$中的所有防御者 $\mathbf{D}$，移位操作后的特征矩阵 $\mathbf{X}$ 变为$\mathbf{SX}$，其中：

$$[\mathbf{SX}]_{if} = \sum_{j=1}^{N} [\mathbf{S}]_{ij} [\mathbf{X}]_j^f = \sum_{j \in \mathcal{N}_i} s_{ij} \mathbf{x}_j^f, \tag{7}$$

这里,邻接矩阵$\mathbf{S}$ 称为图移位算子 （GSO）（Gama 等人，2019年).

4.2.2 图卷积

借助移位操作，我们将图 $\mathcal{G}$ 上经由 K 跳通信交互得到的图卷积 定义为移位后的特征的线性组合（Gama 等人，2019年；Li 等人，2020年）：

$$\mathcal{H}(\mathbf{X}; \mathbf{S}) = \sum_{k=0}^{K} \mathbf{S}^k \mathbf{X} \mathbf{H}_k, \tag{8}$$

其中，$\mathbf{H}_k \in \mathbb{R}^{F \times G}$ 表示移位特征矩阵$\mathbf{S}^k\mathbf{X}$中用于组合防御者 F 特征的系数，F 和 G 分别表示图卷积的输入维度和输出维度。注意，$\mathbf{S}^k\mathbf{X} = \mathbf{S}(\mathbf{S}^{k-1}\mathbf{X})$通过与1跳邻居进行 k 次通信交互计算得出。

4.2.3 图神经网络

将逐点非线性函数 $\sigma:\ \mathbb{R}\ \rightarrow\ \mathbb{R}$作为图卷积（式8）的激活函数，我们将图感知定义为:

$$\mathcal{H}(\mathbf{X};\mathbf{S})=\sigma(\sum_{k=0}^{K}\mathbf{S}^{k}\mathbf{X}\mathbf{H}_{k}). \tag{9}$$

然后，我们通过级联 L 层图感知来定义一个图神经网络模块（公式 9）：

$$\mathbf{X}^{\ell}=\sigma\left[\mathcal{H}^{\ell}(\mathbf{X}^{\ell-1};\mathbf{S})\right]\quad\text{for}\quad\ell=1,\cdots,L, \tag{10}$$

其中，上一层的输出特征 $\ell-1$，$\mathbf{X}^{\ell-1}\in\mathbb{R}^{N\times F^{\ell-1}}$，作为当前层的输入$\ell$，以生成第 l 层的输出特征 $\mathbf{X}^{\ell}$。回顾一下，第一层的输入为$\mathbf{X}^{0}=\mathbf{X}$ （公式 6）。最后一层的输出特征 $\mathbf{X}^{L}\in\mathbb{R}^{N\times G}$，通过 K 跳通信获得，表示防御者团队$\mathbf{D}$交换并融合的信息。

4.2.4 候选匹配

图神经网络的输出表示来自 K 跳通信的融合信息，随后由另一个多层感知机处理，为每个防御者提供一个候选匹配。图 3 展示了 $N_A^f=6$ 时的一个候选匹配实例。给定一名防御者 D_i，我们找出距离最近的 N_A^f 个入侵者，并按顺时针方向将其编号为 1 到 N_A^f。将附近入侵者按顺时针方向编号，主要是为了便于解读网络的特征输出，从而确定哪些入侵者会与哪些防御者匹配。我们也可以按逆时针方向编号，或按任意顺序编号。由于每名防御者都会学习去中心化策略，因此需要根据其局部感知指定一个要捕获的入侵者。入侵者没有全局标识，因此不失一般性，我们只需按顺时针方向为其分配标识符。多层感知机的输出为分配似然 $\mathcal{L}$，表示 N_A^f 个候选入侵者与给定防御者匹配的概率。例如，在图 3 中，如果专家策略（即最大匹配）将第三个入侵者（即 A_3）与 D_i 匹配，那么 D_i 的专家分配似然 L_i^g 将为 $[0.01,0.01,0.95,0.01,0.01,0.01]$。规划模块以最高概率选出候选入侵者 A_j，使匹配对（D_i，A_j）最接近专家策略。值得注意的是，我们的方法实现了去中心化分配策略，因为只交换相邻信息。

4.2.5 置换等价性

值得注意的是，我们提出的基于图神经网络的学习方法具有可扩展性，这是由置换等价性带来的。这意味着，对于一个分散式防御者，它应能够根据由任意数量、未编号的入侵者构成的局部感知来决定行动。图 4 展示了一个周界防御博弈实例，以说明这一特性。图中聚焦于单个防御者，随着时间推移，入侵者逐渐接近周界。不同时间点中，相同的入侵者用相同的颜色表示。请注意，在 $t=2$ 时，一名新的浅蓝色入侵者进入了防御者的视场；在 $t=3$ 时，一名紫色入侵者开始出现。尽管每个时刻检测到的入侵者数量不定，我们的系统仍会为入侵者分配标识符，如图 4 中的蓝色数字所示。我们按顺时针方向为它们编号，但也可以采用任意其他排列方式（例如逆时针），因为图神经网络执行标签无关处理。编号的原因是

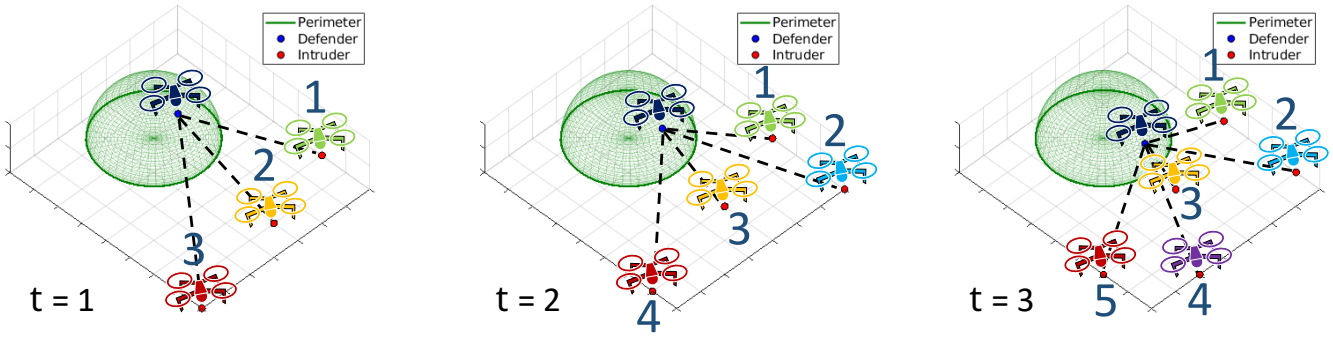


图 4. 不同时刻的周界防御博弈实例。图中重点展示了单个防御者及其局部感知。

指定根据网络输出，哪些入侵者将与哪些防御者匹配。在不失一般性的前提下，我们按顺时针方向分配标识符，但需要指出，这些标识符是任意的，因为它们在不同时间点可能发生变化。例如，黄色入侵者在 $t = 1$ 时的标识符为 2，但在 $t = 2$，3 时变为 3。同样，红色入侵者在 $t = 1$ 时的标识符为 3，但在 $t = 2$ 时变为 4，在 $t = 3$时变为 5。通过这种方式，我们能够适应任意数量的入侵者，因此我们的系统具有置换等价性。

4.3 控制

第 4.2 节的输出将作为输入传递给图 3 中的防御者策略模块。该模块处理所有匹配对（D_i, A_j），并针对每个一对一半球周界防御博弈计算最优突破点（参见第 3.3 节）。防御者策略模块汇总输出位置指令，指令方向指向最优突破点。随后，将由推力和力矩组成、用于在低层级控制机器人的 SO(3) 指令（Mellinger 和 Kumar，2011）传递给防御者团队 **D** 进行控制。防御者-入侵者配对的状态动力学详见（Lee 等人，2020b）。防御者根据指令移动，从而形成感知-动作闭环。值得注意的是，专家分配似然 $\mathcal{L}^g$ 将产生专家动作集 $\mathcal{U}^g$（定义见问题 1）。

4.4 训练流程

为了训练我们提出的网络，我们使用模仿学习来模仿由最大匹配（见第 3 节）给出的专家策略，该策略根据防御者的局部感知 $\{\mathcal{Z}_i\}_{i\in\mathcal{V}}$ 和通信图 $\mathcal{G}$提供最优分配似然 $\mathcal{L}^g$（见第 4.2 节）。训练集 $\mathcal{D}$ 由以下数据集合生成：$\mathcal{D} = \{(\{\mathcal{Z}_i\}_{i\in\mathcal{V}}, \mathcal{G}, \mathcal{L}^g)\}$。我们训练映射 $\mathcal{M}$（定义见问题 1），使 $\mathcal{L}^g$ 与 $\mathcal{L}$之间的交叉熵损失最小。我们证明，训练后的 $\mathcal{M}$ 能够提供$\mathcal{U}$，且与 $\mathcal{U}^g$接近。我们的网络中的可学习参数数量与团队规模 N无关。因此，只要任何规模下的防御者都能基于固定数量智能体的局部感知学习去中心化策略，我们就可以在小规模数据上训练网络，并将模型泛化到大规模场景。

4.4.1 模型架构

我们的模型架构由一个包含 16 和 8 个隐藏层的两层多层感知机构成，用于生成后处理特征向量 $\mathbf{x}_i$；一个包含 32 和 128 个隐藏层的两层图神经网络，用于交换从防御者处收集的信息；以及一个单层多层感知机，用于生成分配似然 $\mathcal{L}$。多层感知机和图神经网络中的各层后均接 ReLU。

参数名称	符号值	
捕获距离	ϵ	0.02
视场	FOV	π
入侵者特征数量	N_A^f	10
防御者特征数量	N_D^f	3
通信范围	r_c	1
默认团队规模	N_{def}	10

表1. 图神经网络实现中的参数设置,

4.4.2 图神经网络详情

在实现图神经网络时，我们通过连接通信范围 $r_c = 1$内的防御者，构建一个 1 跳连通图。鉴于默认半径为 $R = 1$，我们预计 1 跳范围内的三个邻近智能体能够为防御者提供广阔的感知区域。因此，我们假设通信以实时方式进行，且通信速率为 $N_D^f = 3$。每个防御者收集的信息作为输入特征，包括距离最近的 $N_A^f = 10$ 个入侵者位置和距离最近的 $N_D^f = 3$ 个防守者位置。所用参数汇总于表1。

4.4.3 实现细节

实验使用一台 12 核 3.50GHz i9-9920X CPU 和一块 Nvidia GeForce RTX 2080 Ti GPU 进行。我们使用 PyTorch v1.10.1（Paszke 等，2019）实现所提出的网络，并通过 Cuda v10.2 API 加速。我们使用 Adam优化器，动量设为 0.5。采用余弦退火，在 1500 个训练轮次内将学习率从 5×10^{-3} 衰减至 10^{-6}，批量大小为 64。我们选择这些超参数以获得最佳性能。

5 实验

5.1 数据集

我们通过模仿学习评估去中心化网络，其中专家分配策略为最大匹配。周界为半径 R的半球，其定义为 $R = \sqrt{N/N_{def}}$，其中 N 为团队规模， N_{def} 为默认团队规模。由于在大规模情况下（例如 $N > 10$）运行最大匹配的开销非常大，因此我们将默认团队规模设为 $N_{def} = 10$。这样， R 也表示游戏的规模；例如，当 $N = 40$时， R 变为 2，这表示与 $R = 1$时相比，问题设置的规模扩大了一倍。给定团队规模 $N = 10$，我们的实验场地尺寸为 $10 \times 10 \times 1$ 米。在离线阶段，我们随机采样 1000 万个防御者的局部感知样本 $\mathcal{Z}_i$，并找到相应的 $\mathcal{G}$ 和 $\mathcal{L}^g$，以构建数据集；数据集分为训练集（60%）、验证集（20%）和测试集（20%）。

5.2 指标

我们主要关注被捕获入侵者的比例（即捕获数量/入侵者总数）。在小规模场景下（例如 $N \leq 10$），可以运行专家策略（即最大匹配），因此能够直接比较专家策略与我们的策略。在大规模场景下（例如 $N > 10$），运行最大匹配的代价过高。因此，我们将我们的算法与其他基线方法进行比较：贪心算法、随机算法和多层感知机，这些方法将在第5.3节中说明。为了观察小规模和大规模场景下的可扩展性，我们针对每种规模运行五种不同的算法：专家、图神经网络、贪心算法,

随机算法和多层感知机。在所有情况下，我们都会计算绝对准确率，其定义为捕获数量除以团队规模，以验证我们的网络能否泛化到任意团队规模。此外，我们还会计算比较准确率，其定义为图神经网络 的捕获数量与另一种算法的捕获数量之比，以观察比较结果。

5.3 对比算法

在基线算法中，为了公平比较，防御者不会相互传达其“意图”，即哪些入侵者将由哪些相邻防御者捕获，因为图神经网络也不会共享此类信息。对于图神经网络框架，每个防御者会感知附近的入侵者，并通过通信共享所感知入侵者的相对位置，而非“意图”。图神经网络的优势在于能够通过 K 跳通信隐式地学习这些“意图”。因此，去中心化决策（即图神经网络和基线方法的决策）可能会导致多个防御者试图捕获同一个入侵者；而集中式规划器掌握所有防御者的“意图”，能够避免这种情况。

5.3.1 贪心算法

贪心算法可在多项式时间内运行，因此是一个适合与我们基于图神经网络的方法进行比较的候选算法。为公平比较，我们运行一种基于对 D_i 的局部感知 $\mathcal{Z}_i$ 的去中心化贪心算法。我们启用 K 跳邻近通信，使防御者的感知区域得以扩展，仿佛图神经网络的网络通信通道处于激活状态。防御者D_i 基于任何可感知的入侵者 A_j 计算收益 p_{ij} （见第 3.4 节），并贪心地选择使收益 p_{ij} 最小化的分配。

5.3.2 随机算法

随机算法与贪心算法类似，均启用 K 跳邻近通信，以扩展感知范围。在可感知到的入侵者中，防御者 D_i 随机选择一名入侵者来确定分配。

5.3.3 多层感知机

对于多层感知机算法，我们仅通过排除图神经网络模块，单独训练所提框架中的当前多层感知机。通过将我们的图神经网络框架与该算法进行比较，可以观察图神经网络是否带来了改进。

5.4 结果

我们在不同场景中运行周界防御博弈，采用不同的团队规模和初始配置，以评估所学习网络的性能。具体而言，我们在小规模（$N \leq 10$）和大规模（$N > 10$）下进行了实验。图5展示了采用我们提出的网络、不同团队规模下模拟周界防御博弈的俯视图快照。周界、防御者状态、入侵者状态和突破点分别以绿色、蓝色、红色和黄色标记。我们观察到，入侵者试图抵达周界。在确定防御者与入侵者的匹配关系后，入侵者会执行各自的最优策略，向最优突破点移动（见第3.5节）。如果某个入侵者未被任何防御者捕获而成功抵达目标，它便会被消耗，并留下一个标记“入侵”。如果入侵者未能成功并被防御者拦截，则这两个智能体都会被消耗，并留下一个标记“捕获”。入侵者瞄准的周界上的点标记为

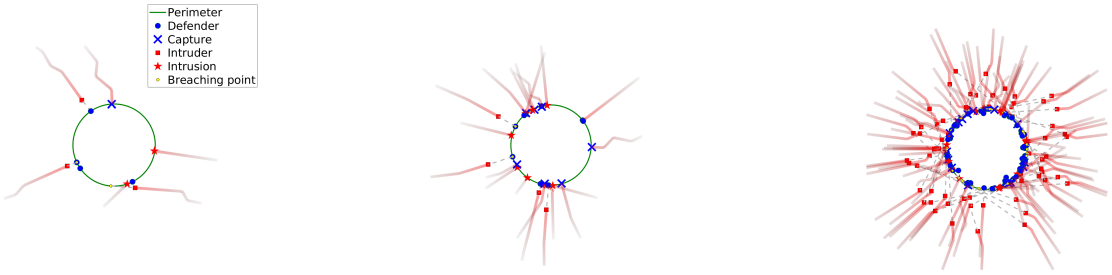


图5a. 6 对 6　　**图5b.** 20 对 20　　**图5c.** 100 对 100

图5. (A)-(C) 采用所提出的图神经网络方法，对三种不同团队规模进行模拟周界防御的俯视快照。

“突破点”。在所有运行中，当所有入侵者都被消耗后，游戏会在终止时间 T_f 结束。更多结果请参阅补充视频。

如第5.1节所述，我们在小规模下运行五种算法：专家、图神经网络、贪心算法、随机算法和多层感知机；在大规模下运行图神经网络、贪心算法、随机算法和多层感知机。例如，图6展示了使用这四种算法模拟的20对20周界防御博弈在终止时间 T_f 时的俯视图快照。四个子图(a)–(d)表明，尽管所有情况下游戏都从相同的初始配置开始，这些算法的表现各不相同。这些算法图神经网络、贪心算法、随机算法和多层感知机分别捕获了12、11、10和7个入侵者。

图7展示了这些方法捕获入侵者比例的总体结果。可以看到，图神经网络 在所有情况下均优于其他基线方法，并且在小规模下表现接近专家。特别是，鉴于我们的默认团队规模 N_{def} 为10，我们提出的算法在团队规模接近 $N = 10$时仍能与专家策略相媲美。

在大规模场景下，图神经网络 的捕获率保持稳定，这表明即使训练是在小规模场景下进行的，训练后的网络也能很好地泛化到大规模场景。贪心算法 的捕获率似乎也保持稳定，但随着团队规模增大，其表现远逊于图神经网络。在小规模场景下，防御者-入侵者匹配对的可选组合很少，因此贪心算法 的表现会与专家算法相近。随着智能体数量增加，可能的匹配数量呈指数级增长，因此问题复杂度大幅提高，贪心算法 的表现也会变差。随机算法 的表现比所有其他算法都差



图6a. 图神经网络　　**图6b.** 贪心算法　　**图6c.** 随机算法　　**图6d.** 多层感知机

图6. (A)-(D) 在终止时间T_f时俯视图下模拟的 20 对 20 周界防御博弈快照，使用四种算法图神经网络、贪心算法、随机算法和多层感知机。使用这些算法的捕获数量分别为 12、11、10 和 7。

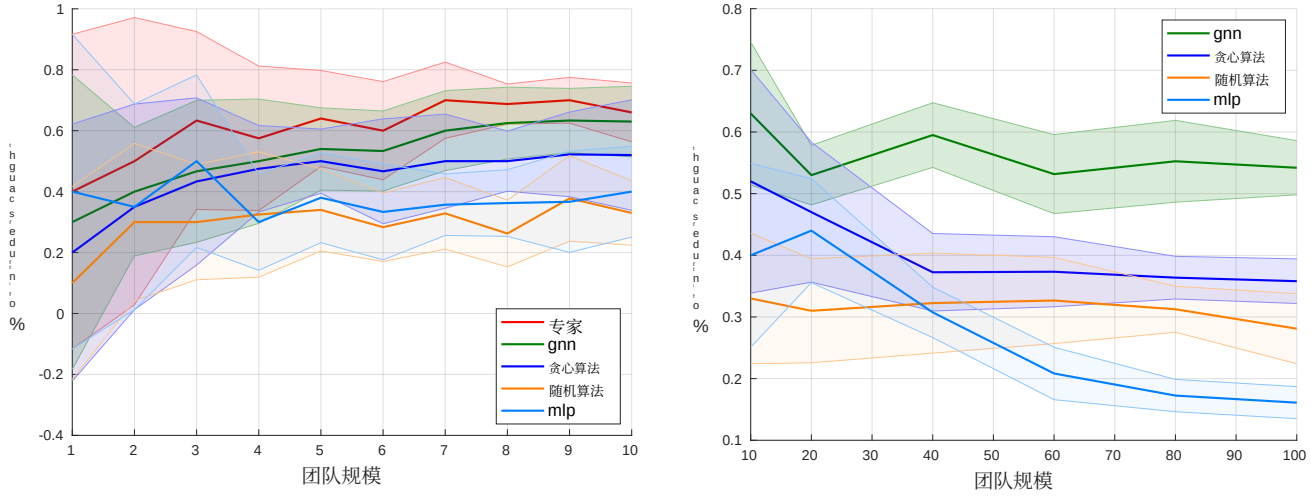


图7. 不同算法在小规模 ($N \leq 10$) 和大规模 ($N > 10$) 场景下捕获入侵者的百分比（10 次试验的平均值和标准差）。

在小规模场景下，其他算法的表现都优于多层感知机；但当团队规模超过 40 时，多层感知机 的表现开始不如随机算法 。这一趋势表明，仅使用多层感知机训练的策略无法扩展到大规模场景。由于训练使用了 10 个智能体，该策略在接近 $N = 10$ 时达到最优，但多层感知机 无法在更大规模下发挥作用，表现甚至不如随机算法 。研究证实，在多层感知机中加入图神经网络能显著提升性能。总体而言，与其他算法相比，图神经网络在大规模场景下的表现优于小规模场景，这验证了图神经网络有助于提升网络的可扩展性。

为了定量评估所提出的方法，我们在表2和表3中报告绝对准确率 和比较准确率 （定义见第 5.2节）。不出所料，当团队规模接近 $N = 10$ 时，绝对准确率达到最大值。除 $N = 2$ 外，绝对准确率的整体数值相当稳定。我们推测，两名防御者之间可能没有共享太多信息，并且根据初始配置，可能根本不存在合理的入侵者。

团队规模	2	4	6	8	10
绝对准确率（图神经网络与 N 的对比）	0.40	0.50	0.53	0.63	0.63
比较准确率（图神经网络 vs. 专家）	0.80	0.87	0.89	0.91	0.95
比较准确率（图神经网络 vs. 贪心算法）	1.14	1.05	1.14	1.25	1.21
比较准确率（图神经网络 vs. 随机算法）	1.33	1.54	1.88	2.38	1.91
比较准确率（图神经网络 vs. 多层感知机）	1.14	1.67	1.60	1.72	1.58

表2. 小规模准确率

团队规模	20	40	60	80	100
绝对准确率（图神经网络 vs. N）	0.53	0.59	0.53	0.55	0.54
比较准确率（图神经网络 vs. 贪心算法）	1.13	1.59	1.42	1.52	1.51
比较准确率（图神经网络 vs. 随机算法）	1.71	1.85	1.63	1.77	1.93
比较准确率（图神经网络 vs. 多层感知机）	1.20	1.94	2.55	3.20	3.37

表3. 大规模下的准确率

图神经网络 与专家 之间的比较准确率表明，随着 N 接近 10，我们训练得到的策略与专家策略越来越接近；我们预计，即使在大规模场景下，图神经网络 的表现也会接近专家。图神经网络 与其他基线方法之间的比较准确率表明，在大规模场景下（$N \geq 40$），我们训练得到的网络表现远胜于基线算法，捕获数量平均多出 1.5 倍。由于随机性的影响，图神经网络 与随机算法 之间的比较准确率在不同团队规模下有些波动，但我们观察到，在小规模和大规模场景下，我们的策略捕获数量平均比随机策略多 1.8 倍。我们观察到，多层感知机 在大规模场景下的表现远逊于其他算法。

根据比较结果，我们证明了所提出的网络虽然在小规模环境下训练，却能够泛化到大规模环境。直观来看，人们可能认为在分散式环境中，贪心算法 会表现最佳，因为每个防御者都会尽力使游戏价值 （定义见公式 5）最小化。然而，我们可以推断，贪心算法 并不了解附近防御者的意图（例如要捕获哪些入侵者），因此无法达到接近集中式专家算法的性能。我们的方法采用图神经网络来交换附近防御者的信息；这些防御者感知各自的局部特征，以规划防御者团队的最终动作。因此，关于附近防御者可能移动方向的隐含信息会传递给每个相邻防御者。由于集中式专家策略了解所有防御者的意图，我们基于图神经网络的策略通过通信通道学习这些意图。防御者团队之间的协作是我们的图神经网络方法优于贪心算法方法的关键。这些结果验证了所实现的图神经网络非常适用于我们的问题，其具备分散式通信的特性，能够捕获邻近交互，并具备迁移能力，从而能够泛化到未见过的场景。

5.5 进一步分析

5.5.1 性能与专家演示数量的关系

为分析算法性能，我们使用数量不同的专家演示（例如，1000 万、100 万、10 万和 1 万）训练了基于图神经网络的架构。团队规模为 $10 \leq N \leq 50$ 时，被捕获的入侵者比例（10 次试验的平均值和标准差）如图 8 所示。该图验证了我们提出的网络能够通过更多演示学习得更好。

5.5.2 性能与周界半径的关系

我们使用不同的周界半径测试了所提出的基于图神经网络的方法。直观而言，在智能体数量固定的情况下，增大半径可能导致防御系统失效。我们将防御者的默认团队规模设为 40，并逐渐增大周界半径，直到被捕获的入侵者比例趋于零。如图9所示，当半径从 100 米增至 800 米时，该比例逐渐下降并趋于零。

5.5.3 性能与感知到的入侵者数量

图10展示了我们的基于图神经网络的方法在感知到不同数量的入侵者（例如， N_A^f）时的性能。我们使用 N_A^f 为 1、3、5 和 10 进行了实验，因为对于大于 10 的数量，没有可用于生成训练数据的真实专家策略。我们观察到，感知到的入侵者特征越多，性能越好。此外，团队规模越大，性能差异往往越小。对于某些团队规模（例如，40），较高的 N_A^f 表现明显更好，但这符合游戏的初始配置。例如，如果初始配置非常稀疏，防御者将受益于较高的 N_A^f，被捕获的入侵者比例也会更高。

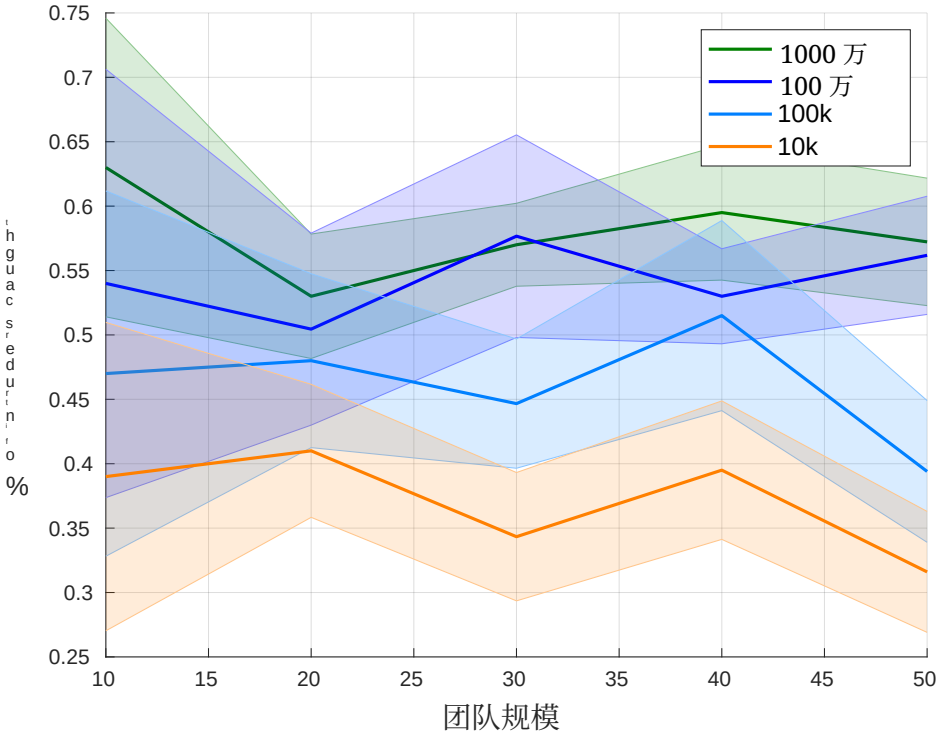


图 8。 不同数量专家演示下的样本效率。

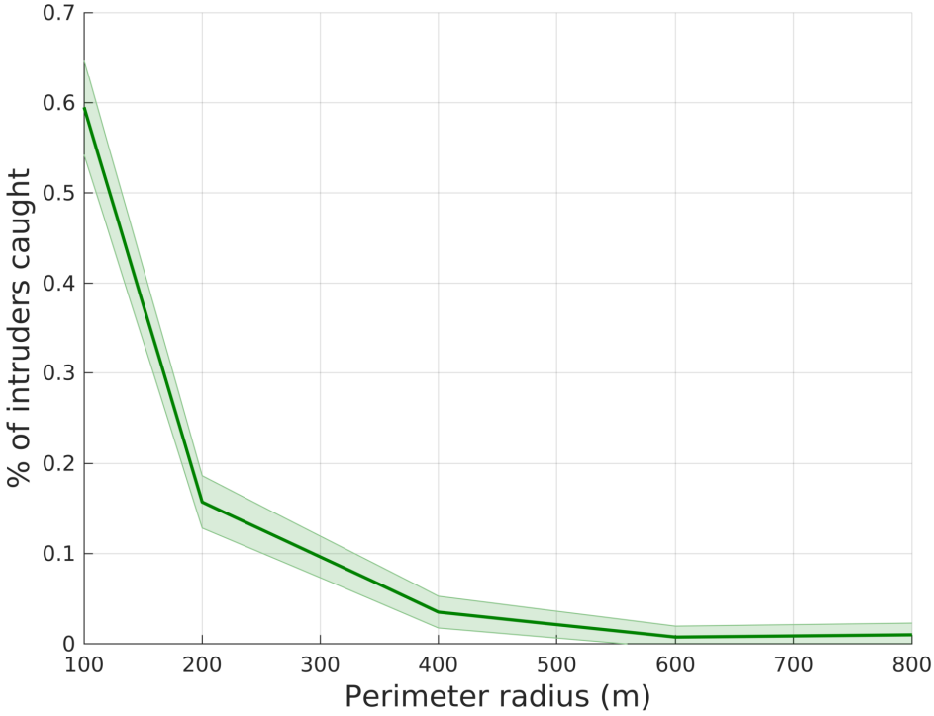
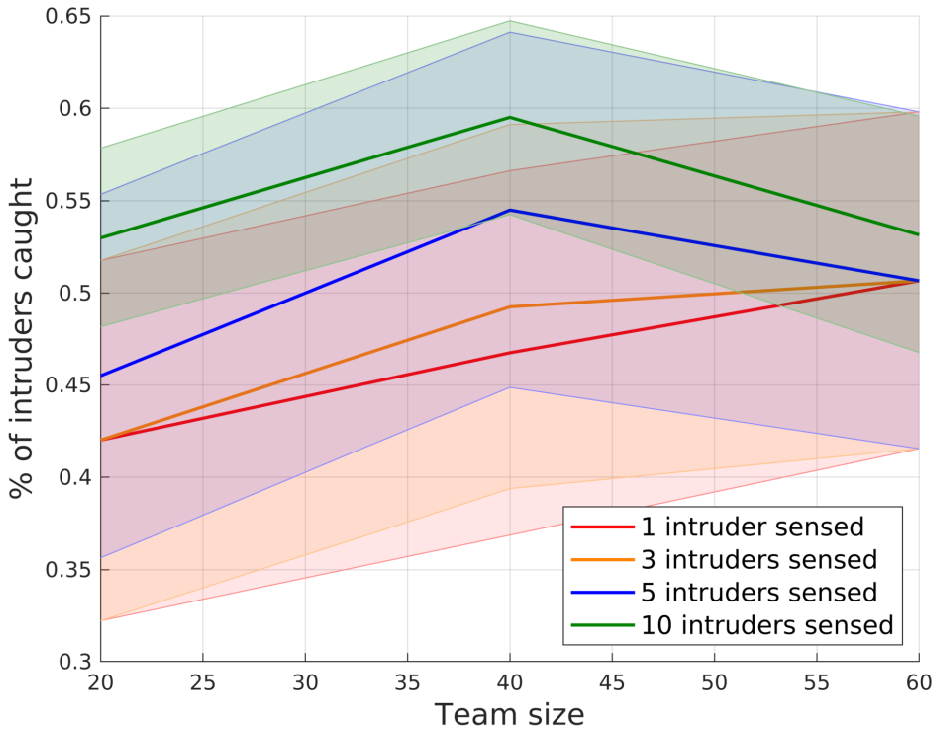


图9. 不同周界半径下被捕获的入侵者百分比。



F图10. 在感知到不同数量入侵者时捕获入侵者的比例　　.

5.6 局限性

由于周界防御是一个相对较新的研究领域，本研究存在一些基本假设所带来的局限性。在问题表述中，我们假设机器人是点粒子。因此，我们假设最优轨迹遵循一阶假设。已有一项初步研究（Lee等人，2021年）致力于弥合点粒子假设与用于一对一半球周界防御的三维机器人之间的差距；我们希望未来将该研究的思路扩展到多智能体周界防御问题中。另一项局限是，在大规模情况下，没有可用的专家策略可与我们提出的方法进行比较。在大规模情况下运行最大匹配算法的代价非常高，因此我们将基于图神经网络的算法与其他基线方法进行比较。尽管所测试算法在不同规模下表现一致，证实了训练后的网络能够泛化到大规模情况，但我们仍希望探索另一种可在大规模情况下用作专家策略的算法，以取代最大匹配算法。一种考虑是采用强化学习，因为这样就能获得算法在大规模情况下的性能。

6 结论

本文提出了一种新颖的框架，利用图神经网络解决去中心化多智能体周界防御问题。我们的学习框架以防御者的局部感知和通信图为输入，并输出动作，以最大限度地增加防御者团队的捕获数量。我们使用基于最大匹配算法的专家策略进行监督，训练深度网络。为验证所提方法，我们在不同团队规模下运行周界防御博弈，并采用五种不同的算法：专家、图神经网络、贪心算法、随机算法和多层感知机。结果表明，我们的基于图神经网络的策略在小规模场景下与专家策略更为接近，且训练后的网络能够泛化到大规模场景。

未来的一项工作是为感知模块实现基于视觉的局部感知，从而放宽对完美状态估计的假设。最终目标是实现具备基于视觉的感知能力、且防御者之间能够通信的多智能体周界防御。另一项未来研究方向是利用强化学习，在多机器人系统中寻找一种集中式专家策略。

数据可用性声明

支持本文结论的原始数据将在不受不当限制的情况下由作者提供。

作者贡献

EL、LZ 和 VK 参与了本研究的构思与设计。EL 和 AR 进行了统计分析。EL 撰写了论文初稿。所有作者均参与了论文修订、阅读并批准了提交版本。

致谢

我们衷心感谢 ARL DCIST CRA 在 W911NF-17-2-0181号资助项目下提供的支持，美国国家科学基金会通过 CCR-2112665、CNS-1446592 和 EEC-1941529 号资助项目提供的支持，美国海军研究办公室通过 N00014-20-1-2822 和 N00014-20-S-B001 号资助项目提供的支持，以及高通研究院、英伟达、洛克希德·马丁和 C-BRIC 的支持。C-BRIC 是由美国国防高级研究计划局共同赞助的半导体研究公司大学联合微电子计划的一部分。

利益冲突

作者声明，本研究开展期间不存在任何可能被视为潜在利益冲突的商业或财务关系。

参考文献

Bajaj, S., Torng, E., Bopardikar, S. D., Von Moll, A., Weintraub, I., Garcia, E., 等（2021）。圆锥形环境的竞争式周界防御。*arXiv*预印本 *arXiv:2110.04667*Baxter, J. L., Burke, E., Garibaldi, J. M., 和 Norman, M.（2007）。多机器人搜索与救援：一种基于势场的方法。载于自主机器人与智能体（Springer），9–16 Chen, A. K., Macharet, D. G., Shishika, D., Pappas, G. J., 和 Kumar, V.（2021）。利用流网络实现最优多机器人周界防御。载于分布式自主机器人系统国际研讨会（Springer），282–293 Chen, M., Zhou, Z., 和 Tomlin, C. J.（2014）。利用低维解和最大匹配求解多人追逃博弈。载于*2014*年美国控制会议（IEEE），1444–1449 Chen, S. W., Nardari, G. V., Lee, E. S., Qu, C., Liu, X., Romero, R. A. F., 等（2020）。Sloam：用于森林清查的语义激光雷达里程计与建图。*IEEE*机器人与自动化快报5，612–619 Gama, F., G. Marques, A., Leus, G., 和 Ribeiro, A.（2019）。用于图上支撑信号的卷积神经网络架构。*IEEE*信号处理汇刊67，1034–1049 Ge, R., Lee, M., Radhakrishnan, V., Zhou, Y., Li, G., 和 Loianno, G.（2022）。面向微型飞行器团队的基于视觉的相对检测与跟踪。*arXiv*预印本 *arXiv:2207.08301*Hsu, C. D., Haile, M. A., 和 Chaudhari, P.（2022）。异构团队周界防御问题的模型。*arXiv*预印本 *arXiv:2208.01430*

Kim, D. K., Liu, M., Riemer, M. D., Sun, C., Abdulhai, M., Habibi, G., 等（2021）。用于在多智能体强化学习中学习如何学习的策略梯度算法。载于国际机器学习会议（PMLR），5541–5550 Lee, E. S., Loianno, G., Jayaraman, D., 和 Kumar, V.（2022a）。通过多视角姿态估计实现基于视觉的周界防御。*arXiv*预印本 *arXiv:2209.12136*Lee, E. S., Loianno, G., Thakur, D., 和 Kumar, V.（2020a）。放射源效应对机器人视觉定位与建图的实验评估与表征。*IEEE*机器人与自动化快报5，3259–3266 Lee, E. S., Shishika, D., 和 Kumar, V.（2020b）。空中防御者与地面入侵者之间的周界防御博弈。载于*2020*年第*59*届*IEEE*决策与控制会议（*CDC*）（IEEE），1530–1536 Lee, E. S., Shishika, D., Loianno, G., 和 Kumar, V.（2021）。利用空中防御者抵御地面入侵者、守卫周界：理论与实践。载于*2021*年*IEEE*安全、防护与救援机器人国际研讨会（*SSRR*）（IEEE），184–189 Lee, E. S., Zhou, L., Ribeiro, A., 和 Kumar, V.（2022b）。利用图神经网络学习周界防御博弈的去中心化策略。*arXiv*预印本 *arXiv:2211.01757*Lee, S., Har, D., 和 Kum, D.（2016）。无人机辅助灾害管理：利用红外摄像机和激光雷达传感器融合寻找受害者。载于*2016*年第三届亚太计算机科学与工程世界大会（*APWC on CSE*）（IEEE），84–89 Lee, Y. 和 Bakolas, E.（2021）。在欧几里得空间中防御攻击者、守护凸目标集。*IEEE*控制系统快报6，1706–1711 Li, Q., Gama, F., Ribeiro, A., 和 Prorok, A.（2020）。用于去中心化多机器人路径规划的图神经网络。载于*2020*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*）（IEEE）Li, Q., Lin, W., Liu, Z., 和 Prorok, A.（2021）。面向大规模多机器人路径规划的消息感知图注意力网络。*IEEE*机器人与自动化快报Liu, X., Prabhu, A., Cladera, F., Miller, I. D., Zhou, L., Taylor, C. J., 等（2022）。多架空中机器人执行的主动度量语义建图。*arXiv*预印本 *arXiv:2209.08465*Macharet, D. G., Chen, A. K., Shishika, D., Pappas, G. J., 和 Kumar, V.（2020）。用于协同多智能体周界防御的自适应分区。载于*IEEE/RSJ*智能机器人与系统国际会议（*IROS*）Mellinger, D. 和 Kumar, V.（2011）。四旋翼飞行器的最小跃度轨迹生成与控制。载于*2011*年*IEEE*机器人与自动化国际会议（IEEE），2520–2525 Miller, I. D., Cladera, F., Cowley, A., Shivakumar, S. S., Lee, E. S., Jarin-Lipschitz, L., 等（2020）。利用多台四足机器人探索矿井隧道。*IEEE*机器人与自动化快报5，2840–2847 Mox, D., Calvo-Fullana, M., Gerasimenko, M., Fink, J., Kumar, V., 和 Ribeiro, A.（2020）。按需部署移动无线网络基础设施。载于*2020*年*IEEE*机器人与自动化国际会议（*ICRA*）（IEEE），7726–7732 Ng, E., Liu, Z., 和 Kennedy III, M.（2022）。需要两人协作：学习规划人机协作搬运。*arXiv*预印本 *arXiv:2209.12890*Nguyen, T., Shivakumar, S. S., Miller, I. D., Keller, J., Lee, E. S., Zhou, A., 等（2019）。Mavnet：一种适用于基于微型飞行器任务的高效语义分割微型网络。*IEEE*机器人与自动化快报4，3908–3915 Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., 等（2019）。PyTorch：一种命令式风格、高性能的深度学习库。神经信息处理系统进展32

Paulos, J., Chen, S. W., Shishika, D., and Kumar, V. (2019年). 通过学习沟通内容实现多智能体策略去中心化. 收录于 *2019*年国际机器人与自动化会议（*ICRA*）（IEEE）, 7990–7996 Ruiz, L., Gama, F., and Ribeiro, A. (2021年). 图神经网络：架构、稳定性与迁移能力. *IEEE*会刊Sharma, V. D., Zhou, L., and Tokekar, P. (2022年). D2coplan：用于多机器人覆盖的可微去中心化规划器. *arXiv*预印本 *arXiv:2209.09292*Shishika, D. and Kumar, V. (2018年). 多人周界防御问题的局部博弈分解. 收录于 *2018*年 *IEEE* 决策与控制会议（*CDC*）（IEEE）, 2093–2100 Shishika, D. and Kumar, V. (2020年). 多智能体周界防御博弈综述. 收录于 国际安全决策与博弈论会议（Springer）, 472–485 Thrun, S., Burgard, W., and Fox, D. (2000). 一种用于移动机器人建图的实时算法及其在多机器人和三维建图中的应用. 收录于 *2000*年 *ICRA* 论文集。千年会议。*IEEE*机器人与自动化国际会议。研讨会论文集（目录编号 *00CH37065*）（IEEE）, 第1卷, 321–328 Tolstaya, E., Gama, F., Paulos, J., Pappas, G., Kumar, V., and Ribeiro, A. (2019年). 利用图神经网络为机器人群学习去中心化控制器. 收录于 *2019*年机器人学习会议（日本大阪：国际机器人研究基金会）Velhal, S., Sundaram, S., and Sundararajan, N. (2022年). 用于周界防御的去中心化多机器人时空多任务分配方法. *IEEE*机器人学汇刊Wang, Z. and Gombolay, M. (2020年). 利用图注意力网络为多机器人协同学习调度策略. *IEEE*机器人与自动化快报5, 4509–4516 Xu, J., D'Antonio, D. S., and Salda˜na, D. (2022年). 模块化多旋翼飞行器：从四旋翼飞行器到全驱动航空器. *arXiv*预印本 *arXiv:2202.00788*Yan, R., Duan, X., Shi, Z., Zhong, Y., and Bullo, F. (2022年). 三维异构多人到达规避微分博弈的基于匹配的捕获策略. *Automatica*140, 110207 Yan, R., Shi, Z., and Zhong, Y. (2019年). 四名等速玩家参与的三维空间到达规避微分博弈屏障构造. 收录于 *2019*年 *IEEE* 第*58*届决策与控制会议（*CDC*）（IEEE）, 4067–4072 Yan, R., Shi, Z., and Zhong, Y. (2020年). 两名防御者和一名攻击者在高维空间中守卫子空间. *IEEE*控制论汇刊Zhou, L., Sharma, V. D., Li, Q., Prorok, A., Ribeiro, A., and Kumar, V. (2021年). 用于多机器人次模动作选择的图神经网络. *arXiv*预印本 *arXiv:2105.08601*