

基于 PIML 的微观人群移动建模仿真与干预决策框架

郭润康^{1,2} 朱正秋^{1,2} 艾川^{1,2} 叶佩军³ 秦龙^{1,2} 尹全军^{1,2} 王飞跃³

摘要 人群移动是影响城市公共安全及应急管理的重要因素, 如何对其进行高精度的建模仿真和有效干预是亟待解决的问题. 为此, 提出一种物理信息机器学习驱动的微观人群移动建模仿真与干预决策框架. 基于平行智能思想, 该框架构建“数据感知-融合建模-动态仿真-智能干预”四层闭环结构, 形成从建模仿真到策略生成、执行及反馈修正的完整链路. 针对人群的移动仿真与引导决策问题, 分别提出基于物理信息时空图卷积网络的导航势能场模型和物理信息多智能体深度确定性策略梯度算法, 有效解决了传统方法中模型准确性较差、仿真与干预孤立以及决策依赖人工经验的问题. 最后, 基于真实数据集开展仿真实验验证了所提框架的有效性.

关键词 微观人群移动; 物理信息机器学习; 导航势能场; 干预决策; 多智能体强化学习

引用格式 郭润康, 朱正秋, 艾川, 叶佩军, 秦龙, 尹全军, 王飞跃. 基于 PIML 的微观人群移动建模仿真与干预决策框架. 自动化学报, 2026, 52(3): 411-429

DOI 10.16383/j.aas.c250312 **CSTR** 32138.14.j.aas.c250312

Microscopic Crowd Movement Modeling, Simulation, and Intervention Decision-making Framework Based on Physics-informed Machine Learning

GUO Run-Kang^{1,2} ZHU Zheng-Qiu^{1,2} AI Chuan^{1,2} YE Pei-Jun³ QIN Long^{1,2}
YIN Quan-Jun^{1,2} WANG Fei-Yue³

Abstract Crowd movement is a critical factor influencing urban public safety and emergency management. How to achieve high-precision modeling, simulation and effective intervention is an urgent issue to be solved. To address these challenges, a physics-informed machine learning-driven framework for microscopic crowd movement modeling, simulation, and intervention decision-making is proposed. Based on the concept of parallel intelligence, the framework establishes a four-layer closed-loop architecture comprising data perception, fusion modeling, dynamic simulation, and intelligent intervention. This architecture forms a complete chain from modeling and simulation to strategy generation, execution, and feedback refinement. For crowd movement simulation and guidance decision-making problems, two novel methodologies are introduced in the framework: A physics-informed spatiotemporal graph convolutional network-based navigation potential field model and a physics-informed multi-agent deep deterministic policy gradient algorithm. These methods effectively resolve issues prevalent in conventional methodologies, namely, the insufficient model accuracy, disjointedness between simulation and intervention, and reliance on human experience for decision-making. Finally, simulation experiments conducted on real-world datasets confirm the effectiveness of the framework.

Keywords microscopic crowd movement; physics-informed machine learning; navigation potential field; intervention decision-making; multi-agent reinforcement learning

Citation Guo Run-Kang, Zhu Zheng-Qiu, Ai Chuan, Ye Pei-Jun, Qin Long, Yin Quan-Jun, Wang Fei-Yue. Microscopic crowd movement modeling, simulation, and intervention decision-making framework based on physics-informed machine learning. *Acta Automatica Sinica*, 2026, 52(3): 411-429

随着全球城市化进程加速, 城市人口持续增长,

给智慧城市建设和带来更大压力的同时, 也对公共安全风险管控提出更高要求. 特别是在商业广场、交通枢纽、文体场馆等大型公共空间中, 人群运动模式复杂多变, 叠加潜在的突发事件风险, 极易引发踩踏事故或造成局部拥堵等群体性安全事件^[1]. 因

收稿日期 2025-07-11 录用日期 2025-11-21
Manuscript received July 11, 2025; accepted November 21, 2025
国家自然科学基金 (72501291), 湖南省自然科学基金 (2025JJ60477) 资助

Supported by National Natural Science Foundation of China (72501291) and Natural Science Foundation of Hunan Province (2025JJ60477)

本文责任编辑 孙仕亮

Recommended by Associate Editor SUN Shi-Liang

1. 国防科技大学数智建模与仿真国家级重点实验室 长沙 410073
2. 国防科技大学系统工程学院 长沙 410073 3. 中国科学院自动

化研究所 北京 100190

1. State Key Laboratory of Digital-Intelligent Modeling and Simulation, National University of Defense Technology, Changsha 410073 2. College of Systems Engineering, National University of Defense Technology, Changsha 410073 3. Institute of Automation, Chinese Academy of Sciences, Beijing 100190

此, 人群移动规律研究已成为城市规划、交通治理与应急管理等领域重点关注方向^[2-3]. 实现高精度的人群动态建模仿真及智能化的干预策略优化, 不仅能有效降低人群安全事故风险, 还能显著提升公共空间的运营效率. 为此, 迫切需要构建一种有效的微观人群移动建模仿真与干预决策框架, 以提升复杂场景下干预策略的精准性和适应性, 实现建模仿真与干预优化的闭环控制.

由于个体多样性、环境复杂性以及行人之间、行人与环境之间复杂的交互作用, 对人群移动行为进行精确刻画具有挑战性. 根据驱动模型的不同, 当前研究主要分为机理驱动、数据驱动和融合驱动三类方法^[4]. 机理驱动方法基于物理原理或专家经验, 通过规则设定或数学模型描述行人移动行为, 如运动学模型^[5-6]、元胞自动机模型^[7-8]、社会力模型^[9-10]、势能场模型^[11-12]等. 该类方法存在规则同质化、缺乏真实数据支撑等问题, 仿真准确性不足; 建模调参复杂, 难以满足实时仿真的计算效率要求. 数据驱动方法通过采集真实运动数据, 利用统计推断、机器学习等技术挖掘潜在行为规律^[13]. 基于深度学习的人群轨迹预测已成为研究热点, 现有工作广泛采用 LSTM (long short-term memory)^[14]、GAN (generative adversarial network)^[15-16]、GCN (graph convolutional network)^[17-19]、Transformer^[20-21] 及扩散模型^[22-23] 等架构提取移动数据中的隐藏时空依赖关系. 然而, 此类方法常将影响行人决策的物理、心理及社会因素编码于神经网络隐状态中, 导致预测结果可解释性差; 长期仿真中误差非线性累积, 模型泛化性与稳定性欠佳^[24]. 因此, 结合深度学习与传统机理模型双重优势的融合驱动方法成为研究重点. 例如, Zhang 等^[25] 开发模仿社会力模型的神经网络来模拟人群轨迹. 但现有方法仍主要依赖神经网络生成机理模型参数, 难以泛化至训练数据分布外的行为, 且多聚焦于短期预测与路径模拟, 难以支撑大规模、长期仿真任务.

尽管部分研究在特定场景中实现了较高精度的人群移动仿真, 但在实际应用中仍存在以下问题: 1) 建模仿真与干预决策脱节. 现有方法通常将两者解耦处理, 干预策略依赖人工经验调参, 缺乏科学性与精准性. 2) 强化学习方法的局限性. 基于 DQN (deep Q-network)^[26-27]、MADDPG (multi-agent deep deterministic policy gradient)^[28-29] 等框架的优化方法虽能减少主观经验依赖, 但缺乏与实时仿真的连接, 智能体决策时缺少数据驱动支撑, 未能充分利用移动数据的时空依赖信息, 无法形成闭环反馈机制, 导致实际干预效果欠佳. 综上, 当前方法在建模层面缺乏数据驱动与物理机理的深度融合;

在仿真层面缺乏时空依赖注入, 难以维持长期稳定性; 在干预层面未形成闭环反馈机制, 策略精准性较差.

针对上述挑战, 受平行智能^[30-31] 思想的启发, 本文提出一种物理信息机器学习 (physics-informed machine learning, PIML) 驱动的人群移动建模仿真与干预决策框架. 该框架采用数据与机理融合建模的方式, 构建“数据感知-融合建模-动态仿真-智能干预”四层结构, 形成从多源数据感知、物理信息融合建模、动态仿真推演到自适应干预引导的闭环链路. 在真实城市开放场景中的实验初步验证了该框架的有效性与可靠性. 该框架可为城市大型公共空间人群管控以及应急疏散提供科学计算支撑, 辅助管理人员进行日常管理与应急决策.

本文的主要贡献如下:

1) 提出一种数据与机理融合驱动的人群移动建模仿真与干预决策框架. 该框架通过数据驱动与物理机理的深度融合, 解决了传统仿真或干预方法缺乏实时数据支撑或过度依赖人工经验/预设规则的问题. 基于此框架构建的应用系统, 能够实现高精度的人群移动仿真推演并生成有效的干预策略.

2) 提出基于物理信息时空图卷积网络 (physics-informed spatiotemporal graph convolutional network, PI-STGCN) 的人群移动导航势能场模型. 在融合建模层, 该模型采用位移联合速度特征的双流时空图架构, 以数据驱动方式建模人群位置变化, 并基于人工势能场, 构建物理信息机器学习驱动的导航势能场, 解决了传统模型在长时间人群移动动态仿真中稳定性和泛化性不足的问题.

3) 提出基于物理信息多智能体深度确定性策略梯度 (physics-informed multi-agent deep deterministic policy gradient, PI-MADDPG) 的人群引导策略生成方法. 该方法在智能干预层融合 PI-STGCN 模型与强化学习, 通过实时感知数据与仿真推演信息自适应地生成引导策略, 并构建“策略生成-平行执行-反馈修正”的闭环机制, 有效克服了传统方法过度依赖人工经验、实际引导效果不佳的问题.

4) 基于 ETH/UCY 真实城市开放场景数据集, 本文对所提框架进行初步验证. 实验结果表明, 所提框架能够有效模拟人群移动模式, 并基于动态仿真数据自适应生成干预策略. 在人群疏散任务中, 相比传统依靠管理人员主观经验和纯数据驱动的干预方法, 平行执行干预策略后疏散效率提升了 10.2%.

本文组织结构如下: 第 1 节介绍 ACP (artificial systems, computational experiments, parallel execution) 方法与平行智能、人群移动模型、多智能体强化学习 (multi-agent reinforcement learning,

MARL) 等相关理论和背景知识; 第 2 节阐述微观人群移动建模仿真与干预决策框架结构与实现细节; 第 3 节基于真实数据集开展微观人群移动仿真与干预实验; 第 4 节对本文进行总结和展望。

1 背景知识

开展本研究涉及多学科理论方法, 本节首先介绍框架设计的总体理论基础——ACP 方法与平行智能, 随后概述框架涉及的人群移动模型、物理信息机器学习、多智能体强化学习的相关背景知识。

1.1 ACP 方法与平行智能

ACP 方法是一套利用计算手段研究复杂系统的理论方法体系^[30-31]。该方法通过构造与实际系统对应的软件定义模型——人工系统, 结合在线学习、离线计算、虚实互动等手段, 将人工系统转化为一个可实验的“社会实验室”。通过计算实验的方式, 人工系统能够为实际系统的运行提供“借鉴”、“预估”和“引导”, 并通过平行执行方式实现对复杂系统的有效管理与控制。基于 ACP 的平行系统解决方案已在多领域得到广泛应用, 并衍生出“平行智能”、“生成式仿真”、“世界模型”等新概念与方法。

“平行智能”旨在构建一个融合数据、知识与行动的闭环, 以深入理解复杂系统^[32]。该方法推动了人工系统从系统分析者升级为数据生成者, 有效克服了传统方法在建模、实验和数据获取方面的瓶颈。因此, 它能够弥合实际系统“小数据”与人工系统“大数据”之间的鸿沟, 并通过在人工系统与实际系统间持续运行“数据-知识-行动”闭环, 实现“深度智能”。“深度智能”具体体现在三个层面: 描述智能使人工系统能够生成模拟实际系统产生的数据; 预测智能探索实际系统在不同情境下的潜在行为, 其生成的大数据与实际系统数据融合, 形成更全面的数据集; 引导智能则驱动人工系统与实际系统共同向期望目标演化。

本文所提框架是平行智能在城市人群移动系统中的具体实践与理论拓展, 框架通过基于物理信息机器学习的人群移动建模和动态仿真, 深度融合物理机理与数据驱动, 实现了高精度的描述智能与预测智能; 同时, 基于 PI-MADDPG 的引导方法自适应生成并执行干预策略, 引导实际人群系统向目标状态演进, 实现了引导智能。

1.2 人群移动模型

行人运动行为可通过不同机理模型进行描述, 例如排队论模型、流体动力学模型、元胞自动机模

型、格子气模型以及社会力模型等^[33]。其中, 排队论模型和流体动力学模型属于宏观模型, 难以刻画人群中个体行为的差异性; 元胞自动机模型和格子气模型属于离散型模型, 通常运用简化规则描述行人在空间中的位置移动; 社会力模型则借鉴了物理学中粒子运动的概念, 将人群的碰撞与拥挤行为类比为粒子间的碰撞和阻塞效应, 能较为真实地刻画人群移动行为特征。

由于场景中存在障碍物, 行人在向目标点运动过程中需进行路径规划, 以适时避开障碍物及其他行人。势能场理论的引入为快速实现全局路径规划提供了有效手段。人工势场法借鉴物理学中的势能场概念, 在环境中构造虚拟力场^[34]。该方法将环境中的关键元素定义为“源”, 表示为 s , 势能场由源激发产生。源可以是障碍物、其他行人、建筑物或诱发行产生特定行为的其他因素。通常, 源被划分为引力源与斥力源。如图 1 所示, 目标点作为引力源吸引行人, 而障碍物则作为斥力源排斥行人; 同时, 行人间既可能相互吸引, 也可能相互排斥。

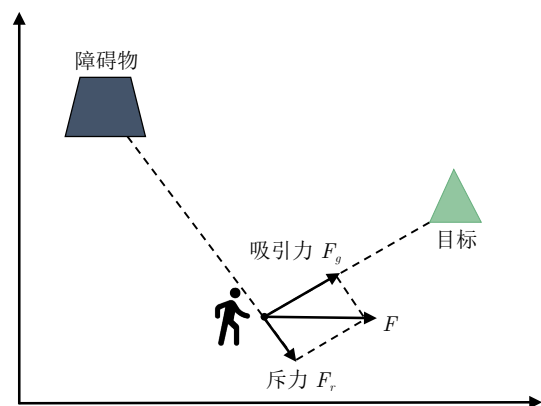


图 1 人工势场原理

Fig.1 The principle of artificial potential field

人工势场模型使用梯度法定义势场力。设由源 s 及其作用域 $D \subset \mathbf{R}^2$ 构成的势能场表示为 $PF_{s, D}$, 定义势能函数为 U 。设行人在场中的位置由二维平面坐标 $P = (x, y)$ 表示, 根据梯度法, 设 $\text{grad}(\cdot)$ 表示求梯度运算, 则行人在位置 P 受到的力 F 计算如下:

$$F = -\text{grad}(U(P)) \quad (1)$$

因此, 势能场可用一个保守力的向量场表示。该向量场将 D 中的任意位置映射到一个二维向量: $PF_{s, D} : D \rightarrow \mathbf{R}^2$ 。

在基于各类人群移动模型的仿真研究中, 多智能体 (agent) 仿真是最为普遍采用的方法^[35]。agent 用于描述具有自适应性和自治能力的硬件、软件或

其他实体^[36]. 在人群移动仿真中, agent 指代用于模拟人类行为与能力的自主实体, 其基本结构如图 2 所示. agent 能够感知环境状态, 根据环境变化自主作出决策, 并执行相应动作以更新其位置. 随着仿真时间的推进, agent 持续更新位置, 仿真系统状态随之动态演变.

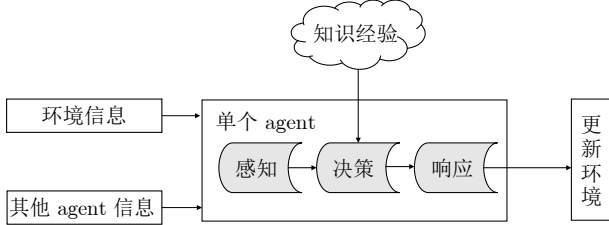


图 2 agent 基本结构
Fig.2 The fundamental architecture of agent

1.3 物理信息机器学习

物理信息机器学习是近年来兴起的一种数据与物理机理相融合的建模方法. 该方法核心在于利用数据驱动的代理模型替代传统机理模型, 并在训练过程中深度融合物理约束, 从而形成数据与机理融合驱动的建模架构^[37]. 其中, 物理信息神经网络 (physics-informed neural networks, PINN)^[38] 是应用最为广泛的一类代理模型.

若给定偏微分方程 $u(x, t)$ 在定义域边界 $\partial\Omega$ 上的边界条件以及在 $t = 0$ 处的初始条件, PINN 能够以无网格方式, 在一组配置点上求解方程. PINN 结构如图 3 所示, 设 $\tilde{u}(x, t; \theta)$ 表示一个神经网络预测器, 其输入为 x 和 t , θ 为网络权重参数. 对任意微分算子 $D(u)$, 微分方程 $D(u) = f$ 可改写为残差形式: $R(u) = f - D(u)$. PINN 的训练目标是寻找一组最优权重参数 θ^* , 最小化损失函数. 其中, 损失函数包含对应方程的初始条件和边界条件损失 MSE_u , 以及在一组有限的配置点上强制执行微分算子, 由方程的残差控制的物理损失 MSE_p . 两种损失的计算方式可表示如下:

$$MSE_u = \frac{1}{N_u} \sum_{i=1}^{N_u} \|\tilde{u}(x_u^i, t_u^i; \theta) - u^i\|^2 \quad (2)$$

$$MSE_p = \frac{1}{N_p} \sum_{i=1}^{N_p} \|R(\tilde{u}(x_p^i, t_p^i))\|^2 \quad (3)$$

其中, $\{x_u^i, t_u^i, u^i\}_{i=1}^{N_u}$ 表示训练中采样的 N_u 个初始条件和边界条件数据点; $\{x_p^i, t_p^i\}_{i=1}^{N_p}$ 为评估残差项 $R(\tilde{u})$ 的 N_p 个配置数据点.

物理信息机器学习融合了数据驱动的学习能力

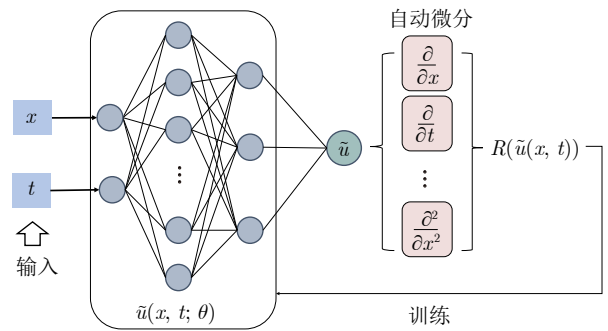


图 3 PINN 结构
Fig.3 The architecture of PINN

与物理机理约束嵌入的优势. 本文所提框架中的“物理信息驱动”, 即利用 PIML 的核心思想与方法, 在人群移动建模中通过 PINN 从实时人群移动数据中挖掘时空规律, 解决行人移动动态仿真和干预策略生成问题.

1.4 多智能体强化学习

本文所提框架实现自适应干预决策主要依赖 MARL 方法, 通过智能体的感知、决策与交互过程, 学习行人引导策略. MARL 是一种针对多智能体系统的决策优化方法, 系统中每个智能体作为独立的学习实体, 根据自身观测信息及其与其他智能体的交互, 学习最优的个体决策策略, 最终实现系统整体性能的最优化^[39-40]. 相较于单智能体强化学习, MARL 能有效处理更复杂的交互模式, 如智能体间的竞争、合作与通信等.

如图 4 所示, 多智能体深度强化学习 (multi-agent deep reinforcement learning, MADRL)^[41-42] 在 MARL 的基础上嵌入深度学习技术, 使智能体具备处理高维、复杂状态与动作空间的能力. 在 MADRL 中, 卷积神经网络、循环神经网络等深度神经网络被作为函数近似器, 用于近似智能体的值函数或策略函数, 从而赋能智能体学习更复杂的状

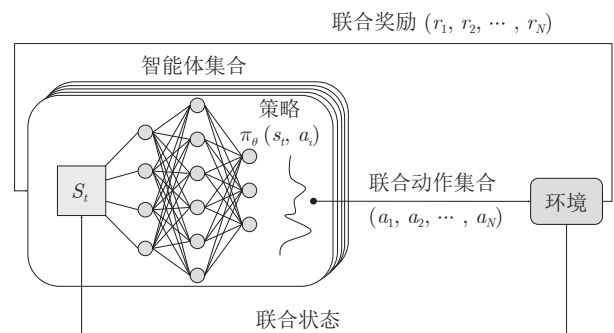


图 4 MADRL 框架
Fig.4 The framework of MADRL

态表征与决策策略。

2 微观人群移动建模仿真与干预决策框架结构与实现

2.1 框架概述

基于“平行智能”思想, 本文创新性地提出一种物理信息机器学习驱动的微观人群移动建模仿真与干预决策框架。如图 5 所示, 框架采用数据与机理融合驱动的四层闭环架构, 由数据感知层、融合建模层、动态仿真层和智能干预层构成。通过各层间的垂直数据流交互与物理信息驱动机制以及跨层级的反馈控制, 该框架能够支撑并实现从多源数据感知、物理信息融合建模、动态仿真推演到自适应引导干预的全流程闭环链路。

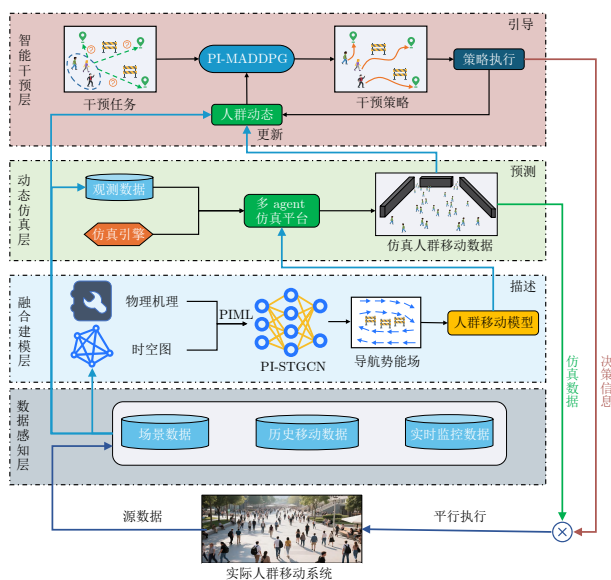


图 5 物理信息机器学习驱动的微观人群移动建模仿真与干预决策框架

Fig. 5 A PIML-driven framework for microscopic crowd movement modeling, simulation, and intervention decision-making

数据感知层作为源数据输入端口, 从实际人群中整合场景数据、历史移动数据与实时监控数据, 经清洗后向融合建模层提供结构化的移动时空数据, 同时向智能干预层传递实时人群动态信息。融合建模层基于 PIML 建模框架, 利用数据感知层输入的时空图数据和行人运动物理机理, 构建 PI-STGCN 模型, 并将输出的人群动态导航势能场传递至动态仿真层, 同时接收仿真层的误差反馈以动态优化仿真模型。动态仿真层基于多智能体仿真技术, 加载导航势能场模型并注入实时人群观测数据,

开展大规模人群移动仿真的计算实验。仿真结果可直接反馈到真实系统, 辅助管理人员对实际人群移动进行控制; 同时向智能干预层提供动态场景推演数据以支撑复杂干预任务的决策。智能干预层以 MADRL 方法为核心, 设计了基于 PI-MADDPG 的人群引导方法, 根据数据感知层的实时信息和仿真推演动态生成干预策略, 策略能够直接或联合仿真输出数据应用于实际人群引导任务。策略执行后, 通过对比分析实际人群系统响应与仿真结果, 迭代优化仿真模型和 PI-MADDPG 策略生成模型, 形成“策略生成-平行执行-反馈修正”的闭环, 最终实现微观人群移动系统的描述、预测以及引导智能。

2.2 方法细节

2.2.1 数据感知层

数据感知层作为真实人群移动信息输入的数据接口, 其核心功能是从实际场景中采集并预处理多模态数据, 为上层模型构建与仿真推演提供结构化数据支撑。在人群移动数据的采集与处理方面, 随着计算机视觉技术的进步, 特别是目标检测、跟踪以及无人机航拍技术的应用, 通过采集人流视频可有效获取行人移动轨迹数据。具体流程为: 首先对视频帧图像进行行人检测, 获取各时刻的行人位置; 随后去除识别结果不稳定的行人, 并在时间维度上进行个体匹配; 最终生成结构化的行人轨迹数据。对于二维平面空间, 在任意时刻 t , 用 (x_t, y_t) 表示行人 i 的位置, 视频检测的初始帧为 t_0 , 帧间隔为 Δt , 则行人 i 的轨迹数据可表示为时空轨迹点序列 $Trajectories$ 。视频记录的数据形成了一段时间内的人群历史移动轨迹; 实时产生的移动数据则随时间的推移不断动态更新, 并补充至历史移动数据中。

$$Trajectories = \begin{bmatrix} t_0, i, x_0, y_0 \\ t_0 + 1 \times \Delta t, i, x_{\Delta t}, y_{\Delta t} \\ \vdots \\ t_0 + k \times \Delta t, i, x_{k \times \Delta t}, y_{k \times \Delta t} \end{bmatrix} \quad (4)$$

场景数据描述了人群移动发生的物理环境约束及其语义信息。针对此类数据, 仿真计算常采用离散化方法, 将环境划分为具有不同属性的面元, 生成场景特征地图。若采用栅格化面元划分方法, 二维平面场景可用矩阵 $Map = [m_{ij}]$ 表示。其中, 元素 m_{ij} 表示坐标 (i, j) 处栅格的环境特征, 既可表示为离散数字以区分不同地物属性, 也可量化为行人通行受阻程度的数值。

2.2.2 融合建模层

融合建模层的核心目标是深度融合数据感知层提供的时空数据与人群移动的物理机理, 以构建高

精度且符合物理规律的人群移动模型. 人群中个体之间复杂的时空交互是影响人群移动演化的关键因素, 图神经网络能够显式建模行人之间的时空依赖关系, 有效捕捉行人之间的动态时空关联, 挖掘其背后隐含的交互规律. 为此, 本文提出一种基于物理信息时空图卷积网络的人群移动导航势能场模型, 模型构建流程如图 6 所示.

首先, 利用数据感知层获取的时空轨迹点序列构建行人时空图, 包含位移时空图和速度时空图. 设每个时间步 t 下的人群位移和速度分布表示为图 G_d^t 和 G_u^t , 其中每个顶点表征一个行人. 本文构建了双流时空图编码结构, 对位移时空图和速度时空图进行独立编码. 编码过程融合了自注意力机制和时空图卷积. 以位移分支为例, G_d^t 首先经自注意力机制计算得到 F_d :

$$F_d = \text{Softmax} \left(\frac{Q_d K_d^T}{\sqrt{d_k}} \right) \times V_d \quad (5)$$

其中, Q_d 、 K_d 、 V_d 是 G_d^t 经线性投影得到的张量; d_k 为查询向量 K_d 的维度大小.

F_d 经 GCN 和 TCN (temporal convolutional network) 进行卷积计算, 进一步提取隐藏时空依赖关系. 基于谱图理论^[43], 图卷积层的传播运算可表示为:

$$H^{l+1} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^{(l)} \right) \quad (6)$$

其中, \tilde{A} 为归一化的邻接矩阵; \tilde{D} 为归一化的度矩阵; H^l 为第 l 层的特征矩阵; $W^{(l)}$ 是滤波器参数矩阵; σ 为激活函数. 时间卷积运算对图卷积输出的特征向量进行时间维度的嵌入, 挖掘序列中的时间依赖性.

为提高模型的准确性及其对真实运动规律的刻

画能力, 本文采用物理信息机器学习注入行人运动的物理机理约束, 对编码得到的位移特征与速度特征进行融合. 特征融合构建物理信息神经网络 $f_p(\theta)$, 学习位移方程 $P(t) = \sum_{k=0}^n a_k t^k$. 物理信息神经网络接收由 MLP (multi-layer perceptron) 解码的原始速度 \hat{v}_0 、位移 $\hat{\mu}_0$ 和时序信息, 输出行人运动方程的系数集合 $\{a_k\}$, 即 $\{a_k\} = f_p(t, \hat{\mu}_0, \hat{v}_0; \theta)$.

PI-STGCN 输出行人未来短期的移动位置分布 $\{\hat{\mu}_t, \hat{\rho}_t, \hat{\sigma}_t\}$, 这三个参数确定了一个二元高斯分布, 分别对应分布的均值, x 方向和 y 方向的相关系数、标准差. 模型训练的目标是使神经网络预测的位置分布尽可能接近真实移动位置分布, 同时确保预测位移满足物理约束. 因此根据行人的位移和速度, 设 J 表示概率分布函数, T_d 表示预测时间步长, 定义似然损失函数 L_d 和物理损失函数 L_p , 计算如下:

$$L_d = - \sum_t^{T_d} \log (J(x_t, y_t | \hat{\mu}_t, \hat{\rho}_t, \hat{\sigma}_t)) \quad (7)$$

$$L_p = \|P(t) - \hat{\mu}_0\|^2 + \left\| \frac{\partial P(t)}{\partial t} - \hat{v}_0 \right\|^2 \quad (8)$$

此外, 为降低预测轨迹分布的发散程度, 引入重构损失 L_c ($L_c = \sum_t^{T_d} \|(x_t, y_t) - \hat{\mu}_t\|^2$). 因此, 模型的总损失函数计算为 $L = L_d + \omega_p L_p + \omega_c L_c$, 其中, ω_p 和 ω_c 为可动态调整的超参数.

通过对 PI-STGCN 输出的移动分布中每个时刻的二元高斯分布进行采样, 可获得离散的移动锚点序列. 为提高仿真计算效率, 使行人能够根据锚点控制的移动方向快速规划路径, 导航势能场模型将连续空间中断的锚点转化为栅格平面上的势能场矩阵, 为上层大规模人群移动仿真计算提供支撑.

首先, 基于锚点控制的移动方向生成势力场.

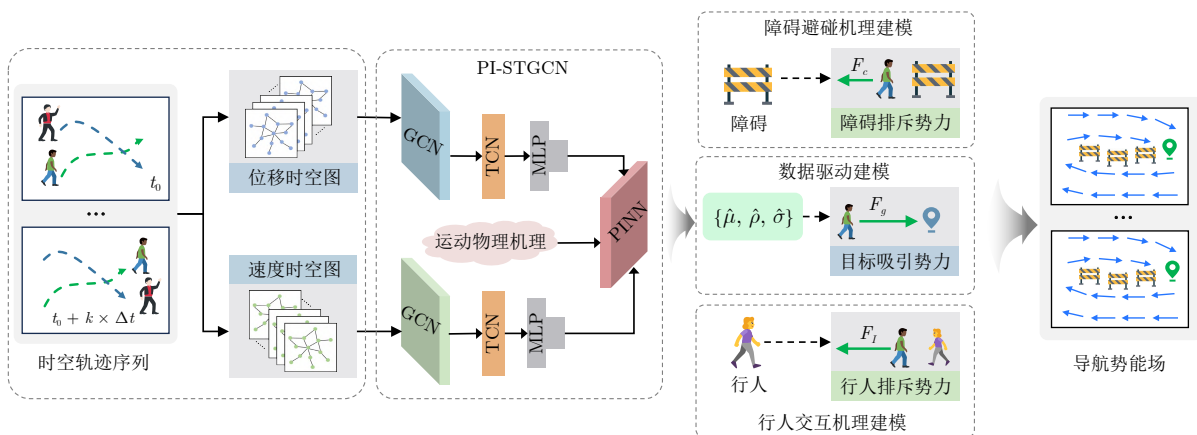


图 6 基于物理信息时空图卷积的人群移动导航势能场模型
 Fig. 6 The navigation potential field model for crowd movement based on physics-informed spatiotemporal graph convolutional network

设经 A^* 算法对采样锚点进行插值后获得的连续锚点路径表示为 $\{g_1, g_2, \dots, g_n\}$, 则行人在位置 g_i 处的势力方向由 g_i 指向 g_{i+1} , 计算如式 (9) 所示.

$$F_{g_i} = k_{g_i} \frac{\langle g_{i+1}^x - g_i^x, g_{i+1}^y - g_i^y \rangle}{|g_{i+1}^x - g_i^x, g_{i+1}^y - g_i^y|} \quad (9)$$

其中, k_{g_i} 是势场强度系数. 对于非锚点路径区域, 势力的方向指向距离其最近的锚点.

在实际的人群移动过程中, 受障碍物和行人交互的影响, 除目标引导的吸引力外, 还需考虑局部排斥力. 障碍物产生的排斥力方向垂直于障碍物边缘, 设排斥力场的作用范围为 δ , λ_o 为表征障碍阻碍强度的比例因子, o 为从障碍表面指向空间某点位置的最短垂直向量, 则障碍物在该位置产生的排斥力 F_c 可表示为:

$$F_c = \begin{cases} \frac{\lambda_o(\delta - |o|) \times o}{(|o| \cdot |o|)}, & 0 < |o| \leq \delta \\ 0, & |o| > \delta \end{cases} \quad (10)$$

同理, 设 ε 为行人排斥力场的作用范围, λ_h 为表征行人避碰强度的比例因子, 可定义行人交互时的排斥力 F_T . 与静态障碍物不同, 由于行人处于移动状态, 行人交互的排斥力场随时间动态变化.

目标吸引力通过 PI-STGCN 以数据驱动方式计算, 而障碍物排斥力与行人间排斥力未采用显式机理建模. 由于现有公开视频数据集中缺乏人群社会关系、心理特征等高质量标注, 周围行人行为及个体心理因素对决策的影响将通过神经网络隐式学习, 并融入目标吸引力的表征中. 因此, 本框架在机理建模层面重点刻画目标引力、障碍物斥力及行人交互力三类基础物理作用, 并假设个体移动由上述势力共同控制, 其合力构成的导航势能场主导行人运动轨迹. 在导航势能场中, 行人具有从高势能区域向低势能区域移动的趋势. 设 t 时刻的势能矩阵用 E_t 表示, 行人在位置 g_i 处所受合力为 $F_G^{g_i}$, g_T 为终点, 由保守力场中力与势能的关系可得:

$$E_t(g_T) = 0 \quad (11)$$

$$E_t(g_i) = \sum_{g_j=g_T}^{g_i} F_G^{g_j} \times |g_{j+1} - g_j| \quad (12)$$

2.2.3 动态仿真层

动态仿真层作为框架的计算核心, 依托融合建模层生成的导航势能场, 基于多智能体仿真技术模拟大规模人群移动的演化过程. 如图 7 所示, 动态仿真层中基于 PI-STGCN 注入实时观测信息, 为智能干预层提供动态场景推演支撑. 在仿真运行过程中, PI-STGCN 作为数据驱动的关键组件, 利用环

境中的实时人群移动数据及历史数据, 持续更新导航势能场. 智能体基于自身观测位置和环境, 依据导航势能场规划移动路径. 通过仿真引擎调度, 智能体执行移动动作并更新环境状态, 仿真进程在循环迭代中持续推进.

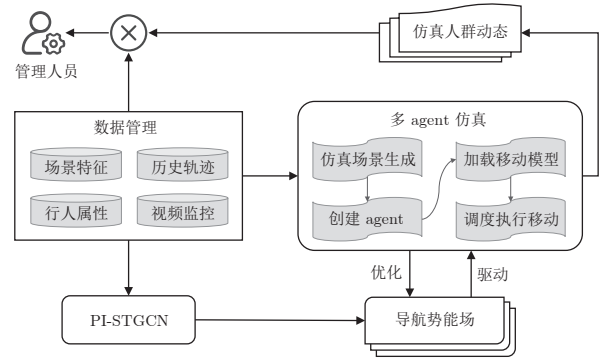


图 7 动态仿真层结构

Fig. 7 Dynamic simulation layer architecture

仿真计算的过程如算法 1 所示, 在每个仿真时间步中, 智能体根据从高势能区向低势能区移动的原则, 从当前栅格未被占用且可通行的相邻栅格中, 优先选择势能最低的栅格作为移动目标. 执行移动动作后, 行人交互产生的排斥力动态更新, 导航势能场也随之更新. 当仿真推进若干步长后, 由于行人的短期移动趋势可能发生变化, 需基于新产生的实时仿真数据更新引导势场. 每经过 T_d 时间步长, 导航势能场基于新产生的历史移动数据由 PI-STGCN 以数据驱动的方式更新.

算法 1. 动态仿真层仿真计算流程

输入. 人群历史移动数据, 实时移动数据, 智能体信息, 场景特征.

输出. 人群移动仿真数据.

- 1) 初始化导航势能场;
- 2) 计算初始时刻的势能矩阵 E_t ;
- 3) **While** 仿真没有停止运行
- 4) $t += \Delta t$
- 5) **For** 每个 agent
- 6) 根据 E_t 选择下一步移动栅格;
- 7) agent 更新当前位置;
- 8) 更新行人斥力场;
- 9) **If** 仿真时间经过一个 T_d 周期
- 10) 利用 PI-STGCN 更新 F_g ;
- 11) 更新导航势能场;
- 12) 存储仿真移动数据.

动态仿真层既向真实系统反馈推演结果, 辅助管理人员定制人群管控预案; 又向智能干预层提供

动态人群分布信息, 支撑实时干预策略生成.

2.2.4 智能干预层

智能干预层是实现动态决策的关键. 该层基于动态仿真层的推演结果与数据感知层的实时监测信息, 依据用户引导任务需求生成人群引导策略, 并通过平行执行的方式验证与优化策略, 最终形成“策略生成-虚实执行-反馈优化”的闭环控制链路. 为将数据驱动的人群移动信息融入引导策略生成模块, 本文提出 PI-MADDPG 算法. 该算法将 PI-STGCN 预测的人群未来状态嵌入强化学习框架, 以提升干预策略的精准性与有效性.

本文将干预任务定义为: 在动态人群环境中, 引导指定行人移动至目标位置. 该任务要求引导路线在确保行人以较短时间抵达目标的同时, 尽可能减少其与周围行人的拥挤. 此任务在常态导航路径规划和应急状态人群疏散中均具有重要应用价值. 针对上述干预任务, 本文提出的 PI-MADDPG 算法流程如图 8 所示. PI-MADDPG 将人群引导问题转化为多智能体协同控制任务. 其中, 智能体定义为引导单元, 其形式可为实际存在的物体 (如引导员、无人车等), 亦可是引导提示信息 (如广播、导航软件或小程序), 本文统称为“引导智能体”.

基于上述干预任务定义, 人群引导策略的制定问题可描述为一个马尔科夫决策过程, 用元组 $\langle S, A, R, P, \rho \rangle$ 表示. 其中, S 为状态集, $A = A_1 \times A_2 \times \dots \times A_N$ 为所有智能体的联合动作集, $R =$

$\{R_1, R_2, \dots, R_N\}$ 为各智能体的奖励函数, $P: S \times A \times S \rightarrow \mathbf{R}$ 为状态转移模型, ρ 为累计折扣奖励的衰减系数. 每个智能体 i 的策略 π_i 是从状态 s_i 到动作 a_i 的确定性映射, 动作 a_i 由个体在 x 方向和 y 方向的速度 $\langle u_x, u_y \rangle$ 表示. 在任意时刻 t , 任意智能体 i 根据当前状态 $s_t^i \in S_i$ 和策略 $\pi_i(a_t^i | s_t^i)$ 选择动作 $a_t^i \in A_i$. 根据状态转移模型 P , 智能体 i 达到下一时刻的状态 $s_{t+1}^i \in S_i$, 同时获得奖励 $r_{i,t} \in R_i$. PI-MADDPG 通过独立训练每个智能体的策略 π_i , 寻找最优策略 π_i^* , 以最大化所有智能体的长期回报 $J_i(\pi_i)$ 期望.

$$J_i(\pi_i) = \mathbb{E} \left[\sum_{t=0}^{\infty} \rho^t r_{i,t} \right] \quad (13)$$

在人群引导任务中, 引导智能体需依据全局态势信息进行决策, 其状态空间应涵盖环境、人群以及其他引导智能体的状态信息. 因此, 每个引导智能体的状态空间定义如下:

$$S_i = \{S_{self}, S_{other}, S_{ped}, S_{obs}\} \quad (14)$$

其中, S_{self} 表示自身状态; S_{other} 是其他引导智能体的状态; S_{ped} 表示人群状态; 本文假设环境状态简化为障碍物的静态分布, 用 S_{obs} 表示. 自身状态包含引导智能体当前的位置坐标及其当前策略维持的速度; 其他引导智能体状态则包含其位置坐标. 设引导智能体总数为 N , 则 S_{self} 和 S_{other} 表示为:

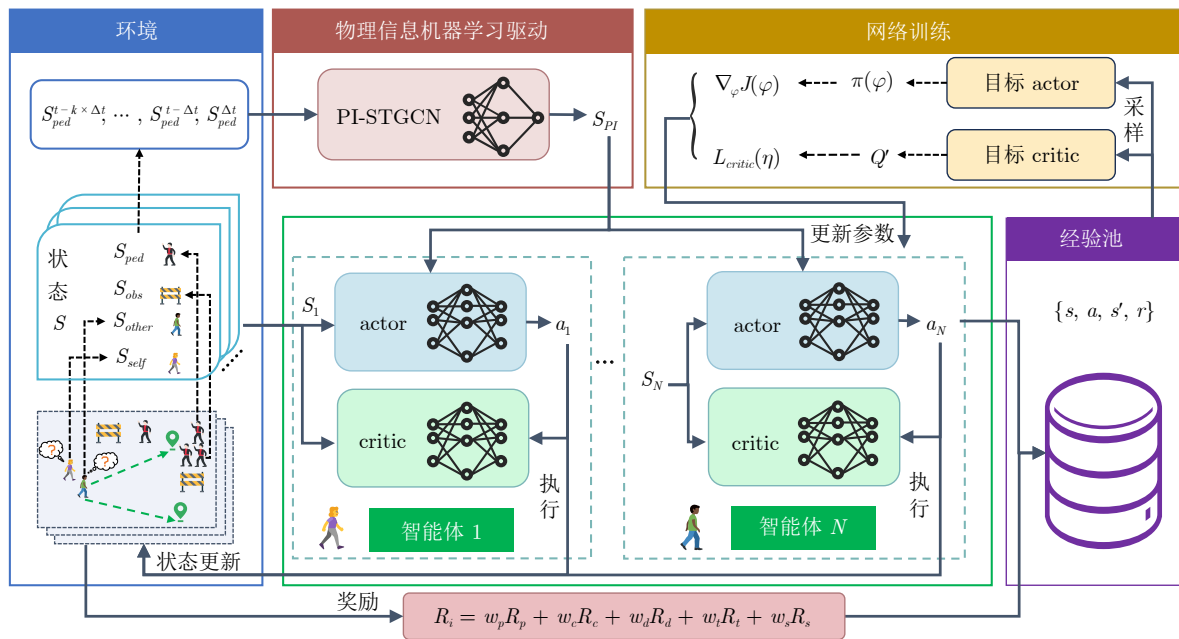


图 8 PI-MADDPG 算法流程

Fig.8 PI-MADDPG algorithm procedure

$$S_{self} = [x^i, y^i, u_x^i, u_y^i] \quad (15)$$

$$S_{other} = [x^j, y^j], \quad \forall j \in N \text{ 且 } j \neq i \quad (16)$$

对于人群状态和环境状态, 采用局部观测的方式, 仅获取引导智能体周围一定距离范围内信息. 设引导智能体的观测范围为半径为 l 的区域 Ω , 障碍物的中心坐标为 (x_o, y_o) , 则

$$S_{obs} = [x_o, y_o], \quad \forall (x_o, y_o) \in \Omega \quad (17)$$

$$S_{ped} = [x_k, y_k], \quad \forall k \in M \text{ 且 } (x_k, y_k) \in \Omega \quad (18)$$

其中, M 为人群数量. 人群是动态变化的, 在 PI-MADDPG 算法中, 人群运动由动态仿真层的输出数据驱动. 除当前观测到的人群位置信息外, actor 网络的状态输入还包含基于 PI-STGCN 计算的未来自位置分布, 记为:

$$S_{PI} = \left[(\hat{\mu}_k, \hat{\rho}_k, \hat{\sigma}_k)_{t+\Delta t}, \dots, (\hat{\mu}_k, \hat{\rho}_k, \hat{\sigma}_k)_{t+p \times \Delta t} \right], \quad \forall k \in M \text{ 且 } (x_k, y_k) \in \Omega \quad (19)$$

其中, $(\hat{\mu}_k, \hat{\rho}_k, \hat{\sigma}_k)_{t+p \times \Delta t}$ 为 PI-STGCN 输出的 $t + p \times \Delta t$ 时刻行人 k 的位置分布; p 为可调节的时间周期长度. 为提高训练稳定性, 状态空间中所有位置坐标在计算前均转换为相对于引导智能体的坐标.

PI-MADDPG 通过智能体与环境交互获取奖励, 并依据最大奖励原则选择动作, 从而获得最优策略. 算法中每个引导智能体的奖励函数设计如下:

$$R_i = w_p R_p + w_c R_c + w_d R_d + w_t R_t + w_s R_s \quad (20)$$

其中, R_p 、 R_c 、 R_d 、 R_t 和 R_s 分别表示进度奖励、碰撞惩罚、人群拥挤度惩罚、时间惩罚和任务完成奖励; w_p 、 w_c 、 w_d 、 w_t 、 w_s 为超参数. 进度奖励驱动被引导的行人快速向目标移动, 对于任意引导智能体 i ,

$$R_p = \alpha \times \Delta d + (1 - \alpha) \times \left(\frac{\| \langle u_x, u_y \rangle \|}{u_{\max}} \right) \times \cos \phi \quad (21)$$

其中, Δd 为当前时刻被引导行人与目标的距离相对于上一时刻的减小量; u_{\max} 为行人速度大小的上限; ϕ 为其当前速度方向与目标方向的夹角; α 为权重系数. 碰撞惩罚根据与引导智能体发生碰撞的物体类型, 按碰撞次数和类型权重计算. 设智能体按策略移动每发生一次碰撞记为 e_c , 则 $R_c = \sum \beta e_c$, 通过差异化权重引导行人优先避免高风险碰撞,

$$\beta = \begin{cases} \beta_{obs}, & \text{障碍} \\ \beta_{ped}, & \text{行人} \end{cases} \quad (22)$$

人群拥挤度惩罚的目的是鼓励引导者在向目标

移动时分散人群, 防止进入人群密集区域造成拥堵, $R_d = \min(M_\varepsilon / M_{\max}, 1)$, M_{\max} 是行人半径 ε 范围内可容纳的最大行人数量; M_ε 为实际的行人数量. 时间惩罚对到达目标所花费的移动步数进行惩罚. 为提升引导任务成功率, 设置任务完成奖励, 假设当行人距离目标小于阈值 ζ 时, 判定任务完成并给予奖励 R_s , R_s 为常数.

PI-MADDPG 算法采用集中式训练、分布式执行的框架. 测试阶段, 算法直接利用训练好的 actor 网络, 根据全局状态 S 输出所有引导智能体的联合动作 A . 动作执行后, 环境更新至新的全局状态 S' , 从而完成决策-执行循环, 其训练过程详见附录 A.

3 案例实验

为验证框架中模型与算法的有效性, 本文基于城市开放街区的真实数据集进行案例研究, 开展仿真实验并设计干预策略, 深入探索复杂环境中的人群移动行为模式及其引导决策方法.

3.1 数据集与实验设置

实验在两个真实人群移动数据集 ETH^[44] 和 UCY^[45] 上进行. 其中, ETH 数据集包含 eth 和 hotel 两个场景; UCY 数据集包含 univ、zara1 和 zara2 三个场景. 数据集涵盖复杂现实环境中逾 1 500 名行人的移动轨迹, 各场景在人群规模、数据分布及行人交互方面均存在差异. 基于该数据集, 本文构建五个场景对应的栅格化仿真环境, 并从各场景中提取行人信息构建 agent 进行仿真实验, 各场景 agent 信息详见表 1. 仿真步长设置为 0.4 s, 与视频数据帧间隔保持一致. 栅格边长取值为 0.4 m, 其大小对应一个行人保持常规站立所需的空间. 通过开展预实验对系统模型进行迭代优化, 仿真模型参数详见附录 B 的表 B1.

表 1 人群移动仿真实验 agent 信息
Table 1 Agent information for crowd movement simulation experiments

场景	eth	hotel	univ	zara1	zara2
agent 数量	255	112	364	139	185
平均速度 (m/s)	1.63	1.01	0.89	1.24	1.19

3.2 人群移动仿真实验结果与分析

实验通过比较所有行人的实际位移/速度分布与仿真位移/速度分布验证仿真结果的准确性. 如图 9 所示, 在多数场景中, 仿真结果的位移与速度分布与实际移动的分布高度相似. 在部分场景下,

受人群自身特性及环境影响, 位移与速度分布呈现双峰特征. 该结果说明基于 PI-STGCN 的导航势能场模型成功捕获了行人移动数据中的复杂时间依赖性, 使得仿真生成的人群移动距离与速度分布能够复现出该双峰特性. 在人群流动的宏观层面, 仿真得到的人群流量变化与实际流量对比结果如图 10 所示. 在五个场景下, 仿真数据较好地复现出人群流量变化的整体趋势, 进一步表明模型能够有效建模复杂的人群流动特征.

以 eth 场景为例, 图 11 展示了部分行人仿真轨

迹与真实轨迹的对比情况. 该场景中行人的移动轨迹整体近似直线, 仿真轨迹与真实轨迹较为接近. 当行人运动至将发生局部偏转或波动时, 动态仿真层通过 PI-STGCN 的实时数据驱动更新导航势能场, 仿真智能体依据更新后的势能场调整移动方向, 从而较好地复现行人的局部偏转行为. 本文所提方法融合真实系统的数据与物理机理, 增强了仿真结果的稳定性和可解释性, 能够有效适应长时间的人群移动仿真任务.

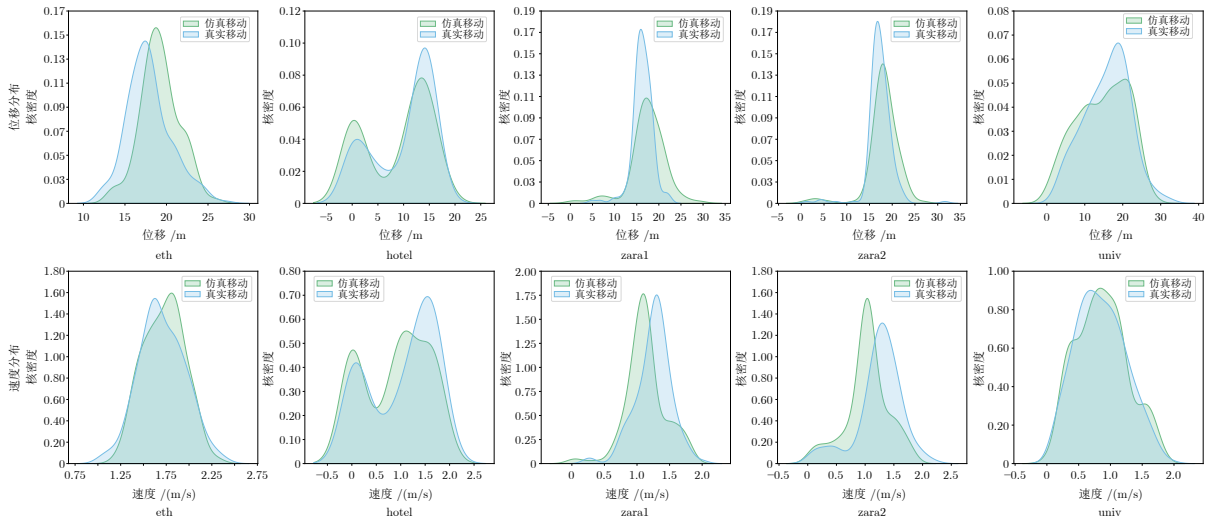


图 9 仿真移动距离和平均速度的核密度分布与真实数据对比

Fig.9 Comparison of simulated movement distance and average velocities kernel density distributions with real data

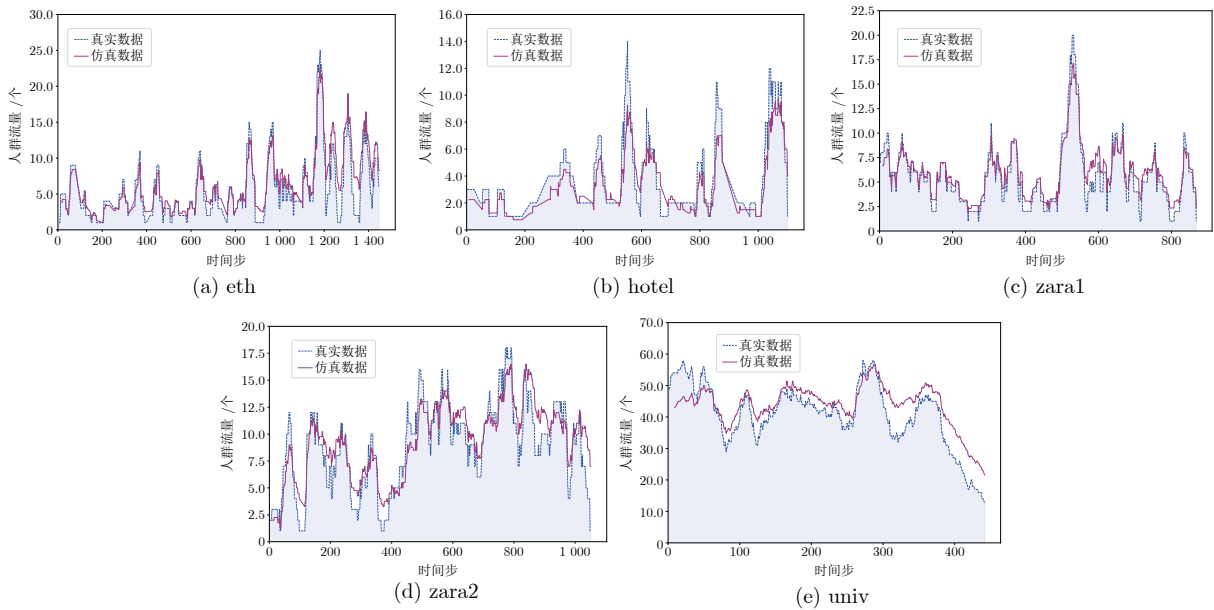


图 10 五个场景仿真人群流量变化与真实流量变化情况比较

Fig.10 Comparison of simulated and real crowd flow variations across five scenarios

3.3 人群移动干预引导实验

现实场景中, 往往需要通过有效的引导策略优化人群移动行为, 以解决拥堵、滞留等问题. 与图 9 所示的仿真案例相对应, 本文选择 eth 场景验证 PI-MADDPG 引导策略优化算法的有效性. 该场景在障碍物布局的影响下, 形成开放单通道结构: 通道一端出口较为开阔, 另一端则相对狭窄, 整体未被分隔为多个独立通行区域. 行人在通道中的运动方式与仿真设定保持一致, 以不同方向在出入口之间移动. 实验设置包括单引导目标和多引导目标两类干预任务, 并对引导策略的优化结果进行分析与评估, 实验参数配置见附录 B 的表 B2.

两类引导任务的模型训练过程如图 12 和图 13 所示, 图中展示了平均奖励的收敛趋势, 以及引导路径的碰撞次数和任务成功率随训练轮次的变化. 由图 12 可见, 随着训练轮次增加, 平均奖励水平总体呈上升趋势. 尽管单轮次奖励呈现显著波动, 但经过充分训练后平均奖励趋于稳定. 具体而言, 单目标引导任务的收敛速度与稳定性均优于多目标引导任务: 前者在约 1 000 轮训练后即收敛, 而在目标数量为 3 的多目标任务中, 训练需至 4 000 轮左右才趋于收敛, 且仍存在明显波动. 在任务成功率和引导智能体与行人/障碍物的碰撞方面, 随着训练轮次增加, 成功率逐渐提升, 碰撞次数波动下降. 单目标任务成功率相比多目标任务收敛趋势更

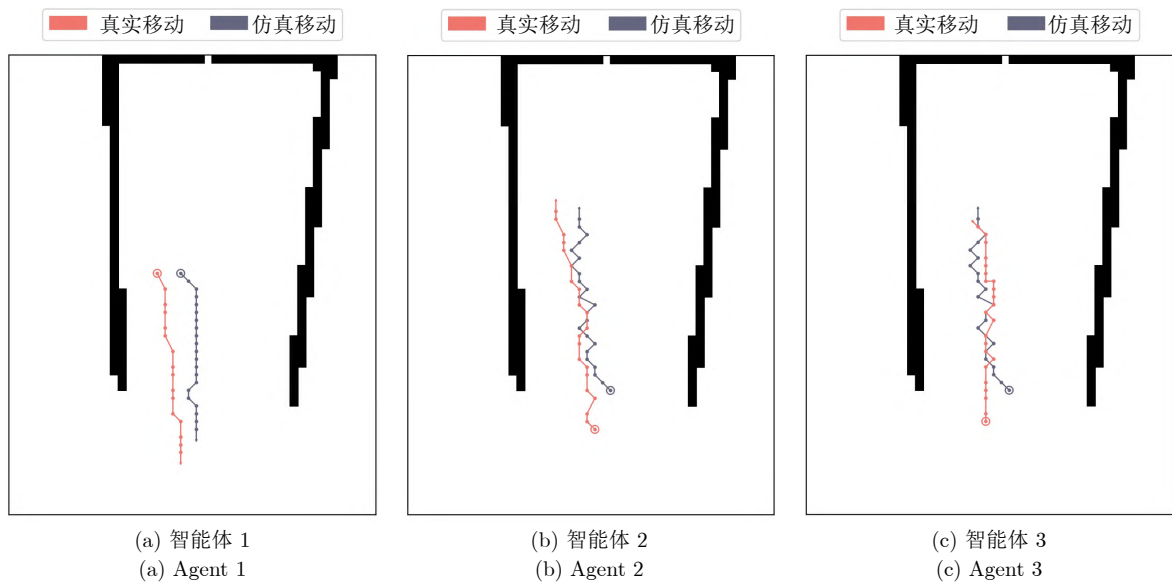


图 11 eth 场景中 3 个行人的仿真轨迹与真实轨迹比较

Fig.11 Comparison of simulated and real trajectories for three pedestrians in the eth scenario

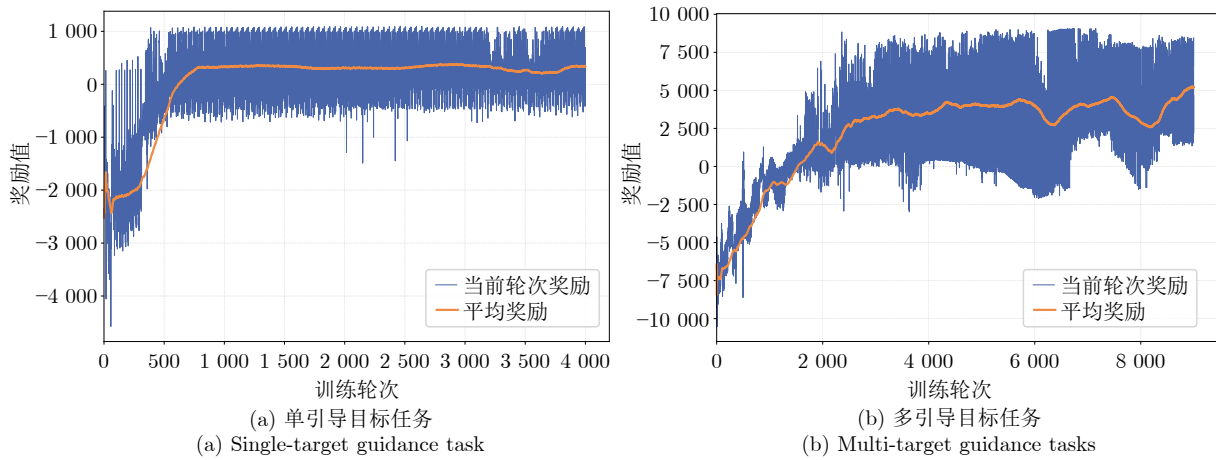


图 12 单引导目标和多引导目标任务平均奖励训练曲线

Fig.12 Average reward training curves for single-target and multi-target guidance tasks

著,原因是目标数量增加导致引导智能体选择目标的复杂性提升,策略优化难度增加.

随机选取 3 个位置点作为被引导行人的初始位置,行人在引导干预下朝向目标移动的瞬时状态及其完整轨迹如图 14 所示,箭头表示人群当前的移动方向.在单目标引导任务中,智能体的移动路径并非遵循直线最短路径指向目标.由于人群中行人的运动以及智能体间的相互避让,在实时的人群移动信息驱动下,智能体在向目标移动过程中能够自

适应调整方向和速度,以避开人群密集区域.多目标引导任务中,智能体的移动策略同样不以直线最短为目标,其移动目标会依据人群动态及智能体的实时位置动态调整.在初始阶段,3 名引导智能体均倾向于选择中间路线.随后,为缓解因智能体 1 和智能体 2 的路径重叠引发的局部拥堵,智能体 3 主动进行避让,其运动轨迹逐渐向右侧区域偏移.总之,PI-MADDPG 算法能够成功规划有效的引导策略,引导行人在动态人群场景中完成不同类型目标

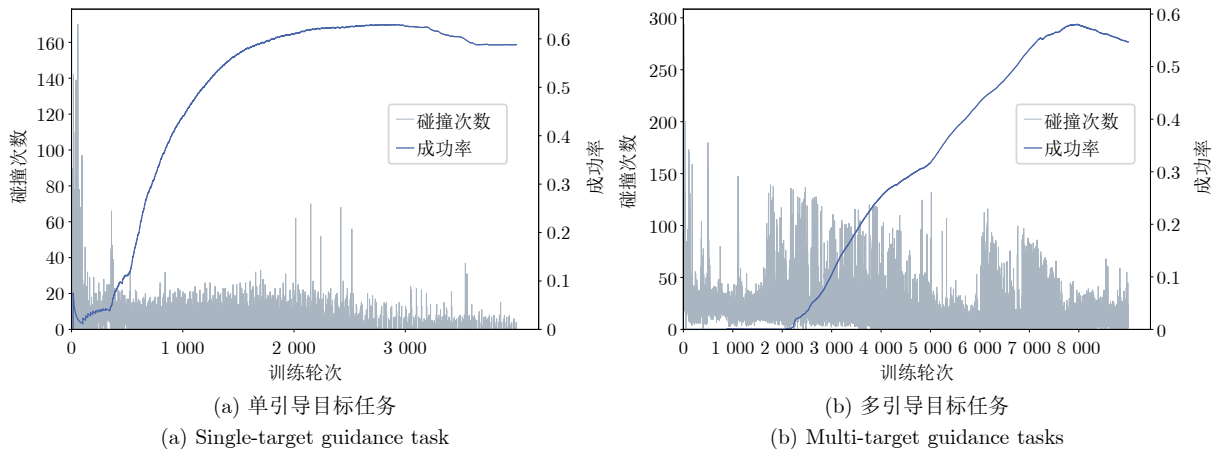


图 13 单引导目标和多引导目标任务成功率与碰撞次数训练曲线

Fig.13 Training curves of task success rate and collision frequency for single-target and multi-target guidance tasks

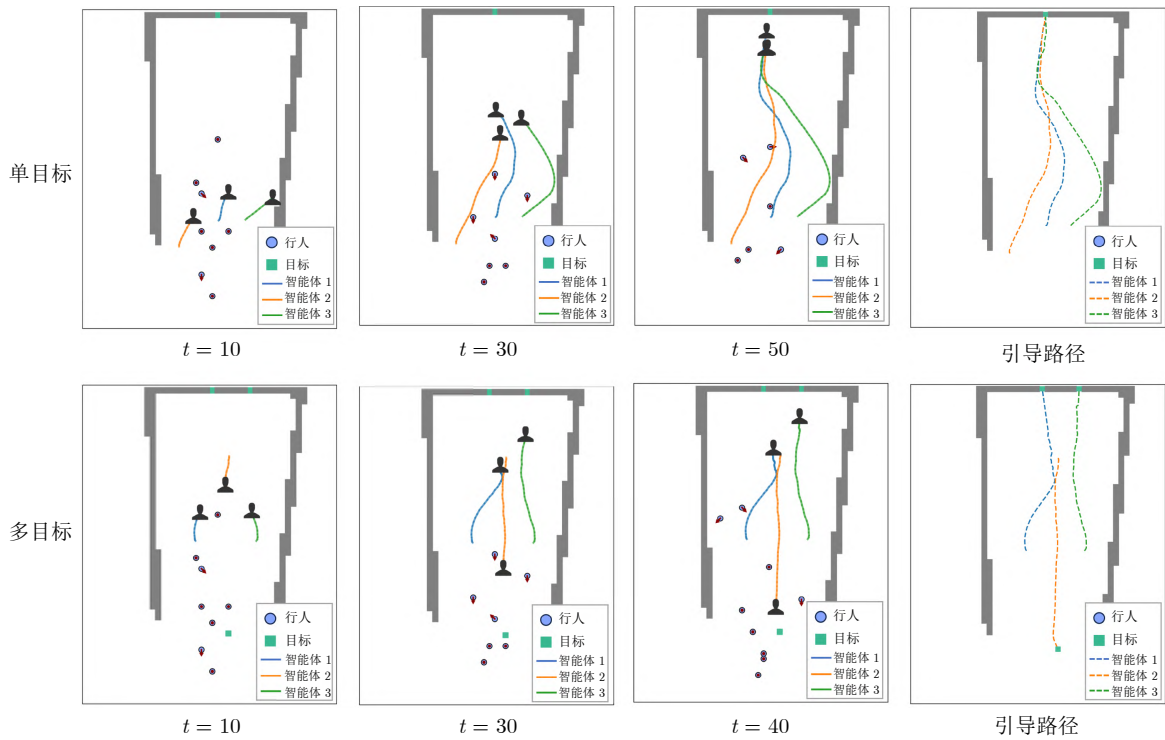


图 14 行人在引导干预下的移动瞬时状态及完整轨迹

Fig.14 Instantaneous states and complete trajectories of pedestrian movement under guidance intervention

的引导任务。

为深入探究 PI-STGCN 对行人交互的影响, 针对多目标引导任务开展消融实验。在消融设置中, 引导算法直接利用 MADDPG 根据当前环境学习策略, 未注入数据驱动的人群移动未来位置分布信息。实验记录所有行人在向目标移动过程中周围人群密度的变化情况, 如图 15 所示。结果显示, 相较于仅依据人群当前状态进行引导的传统 MADDPG 方法, PI-MADDPG 智能体移动时其周围的人群密度更低。算法利用 PI-STGCN 注入人群移动的未来位置分布信息, 从而进一步优化引导策略。

3.4 平行执行

基于“平行智能”的思想, 将框架中的人群移动仿真和干预决策计算实验结果在真实人群系统中平行执行, 以实现人群管理的虚实闭环控制。以人群疏散任务为例, 本节基于 eth 场景数据构建了人群通行与疏散的仿真环境 (见图 16)。为模拟现实场景中通道内的常见拥堵现象, 该实验设定人群持续从一侧开阔入口涌入, 并由另一侧狭窄出口离开。行人的移动行为以数据集中提供的真实轨迹作为输入。实验在动态仿真层中引入人流强度参数 γ , 用以设置单位时间内进入通道的行人数量。不同大小的人流强度将导致不同的通行拥挤程度和疏散速度。由于通道容量的限制, 实验中设置入口在开放一段时间后关闭 (即进入限流阶段), 人群停止进入。当监测到人群已发生拥堵时, 智能干预层启动平行执行: 一方面增加出口; 另一方面通过 PI-MADDPG 生成干预策略引导人群疏散, 并以迭代方式采集干预后的人群移动信息反馈至动态仿真层和智能干预层。

为探究物理信息机器学习驱动的引导策略相较于传统引导方法的优势, 将 PI-MADDPG 与基于最短路径规划、用户草图^[46]以及缺乏数据驱动预测

的 MADDPG 基线方法进行对比 (见附录 B)。

设干预前通道出口数量为 1, 当人群发生拥堵后, 出口数量增加至 3, 且假设所有行人均遵守引导策略。不同人流强度以及拥堵发生后采取引导策略下通道内的人群密度变化如图 17 所示。当 γ 较小时, 行人开始进入通道后, 人群密度经历短暂上升并在波动中趋于相对稳定。此时, 通道内离开与进入的行人数量基本平衡, 在人群密度未超过临界值的情况下可维持稳定流动, 形成稳态。当仿真时间到达 1 000 步后, 因限流措施生效且无新增人群, 密度随人群离开而迅速下降。当 $\gamma \geq 1.0$ 时, 稳态逐渐消失, 限流前人群密度持续增长, 其增长模式呈现先快后缓的特征, 表明单位时间内进入通道的行人数量已超过离开的数量。此时, 人群虽仍能缓慢移动至出口, 但密度持续上升。当密度达到通道容量极限时, 将引发拥堵阻塞。如图 17 所示, 当 $\gamma \geq 1.5$ 时, 在限流阶段前均已发生拥堵, 且 γ 越大拥堵时通道内人群密度越高。以 $\gamma = 1.6$ 条件下的人群疏散为例, 在无引导措施时, 第 250 仿真步长即开始拥堵, 限流前人群密度最高时升至接近 4 人/平方米。干预措施启动后, 在所有引导方法下人群密度的上升速率均出现下降。然而, 基于草图与最短路径的方法在人群持续进入的情况下, 密度仍以较快速度上升, 未能重新恢复稳态。传统 MADDPG 算法在初期 (仿真步长数小于 500) 表现出较好的干预效果, 人群流动重新趋于稳态; 随着行人持续涌入, 密度再次上升, 但其上升速率明显低于前两者。相比之下, 采用 PI-MADDPG 进行引导干预后, 人群流动恢复并维持在稳态, 拥堵现象基本消除。在疏散效率方面, PI-MADDPG 优于所有基线方法, 其平均通行时间较传统 MADDPG 缩短约 10.2%。原因是基于规则或经验的方法 (如最短路径和草图引导) 路径固定, 缺乏对动态人群的适应能力; 传统 MADDPG 虽具备一定基于实时数据感知的决策能

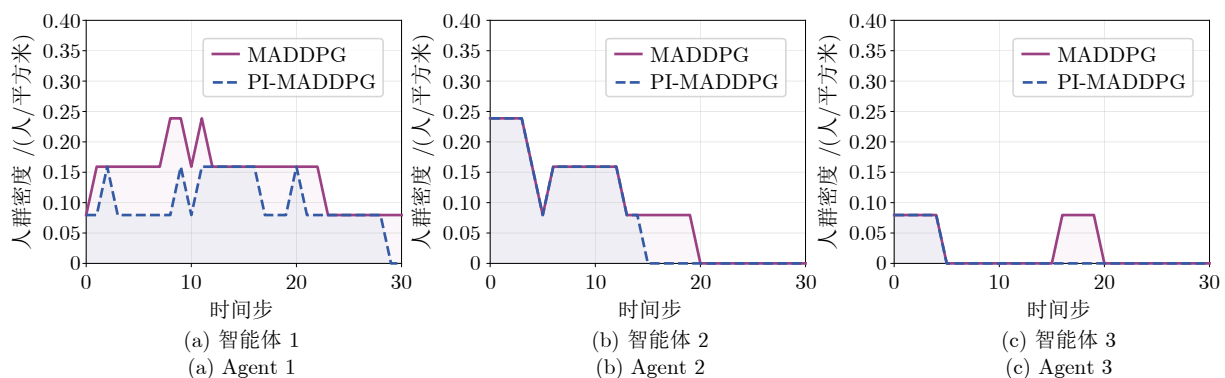


图 15 行人移动过程中周围人群密度变化

Fig. 15 Variation of local crowd density during pedestrian movement

力,但未引入数据驱动的人群未来移动分布信息,导致引导策略在中期以后逐渐失效. PI-MADDPG 通过融合物理信息与数据驱动的预测,能够

更准确地建模人群移动趋势,从而实现持续有效的动态引导,提升整体疏散效率.

借助平行执行, PI-MADDPG 利用多智能体强化学习制定疏散策略,有效解决了潜在拥堵问题,优化了真实人群流动. 在多轮的平行执行中,通过实时反馈动态调整仿真模型参数与策略参数,不断优化决策,能够为人群疏散任务提供更精准的管控方案.

3.5 复杂度与实时性分析

在平行执行案例实验中,框架各功能层次协同运行,形成完整的闭环计算流程. 为进一步评估框架的可行性与效率,本节从理论与实验两方面对各层次的计算复杂度及系统整体实时性进行分析. 在数据感知层,可利用成熟的人群目标检测技术高效识别并获取行人移动信息. 因此,为简化流程,仿真实验中直接使用仿真器输出的轨迹数据作为感知结果. 融合建模层主要包括时空图卷积与导航势能场的计算,设时空图特征维度为 C_1 , 行人数量为 N , 基于预训练模型从历史轨迹中计算导航势能场的复杂度为 $O(N \times C_1^2 + N \log N)$; 动态仿真层需周期性地更新智能体位置与全局导航势能场, 设总仿真步数为 S , 势能场更新周期为 T_d , 则仿真计算

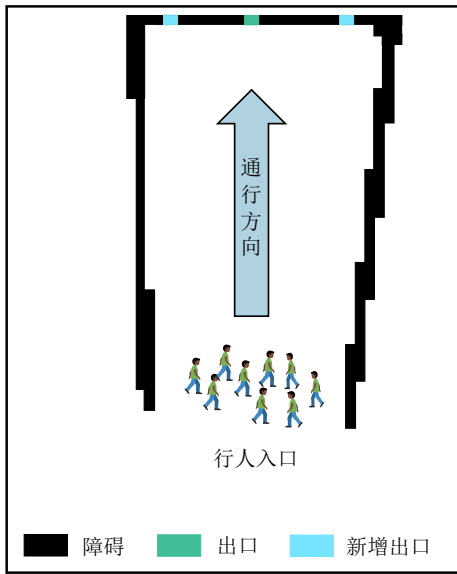


图 16 单向通道人群流动场景示意

Fig. 16 Illustration of a unidirectional passage crowd flow scenario

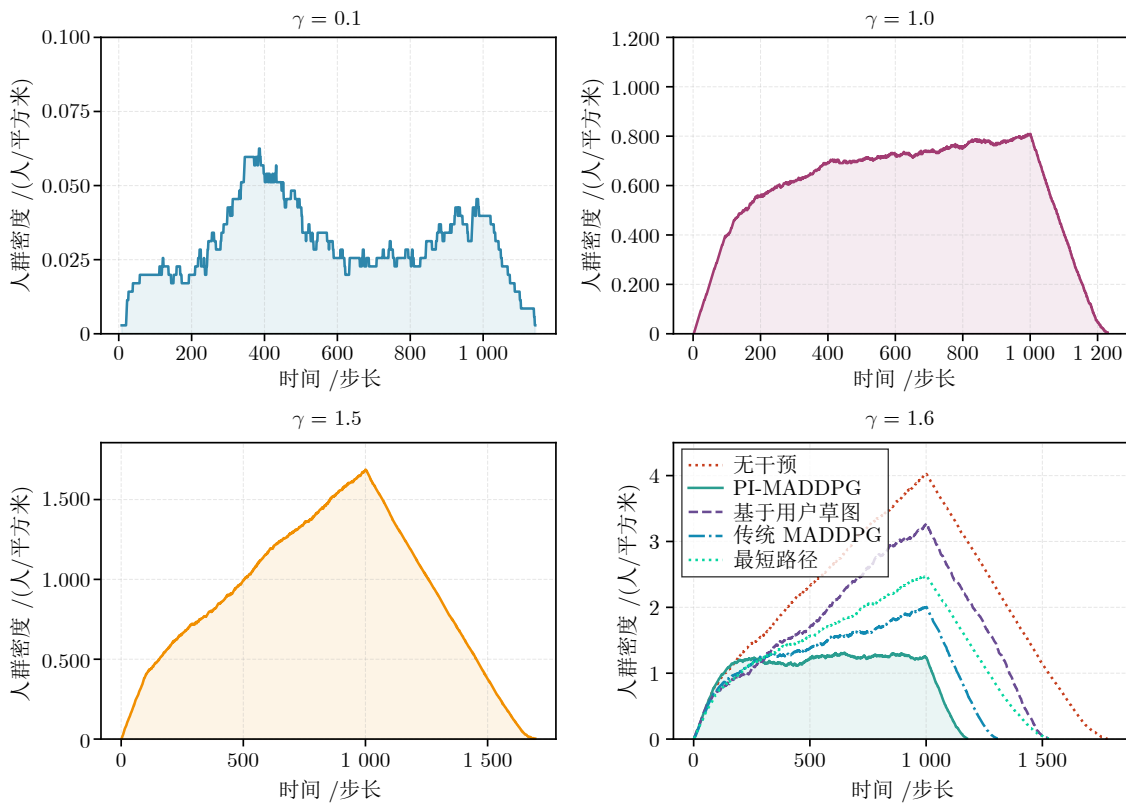


图 17 疏散过程中通道内人群密度变化曲线

Fig. 17 Crowd density variation curve in the passage during the evacuation process

的复杂度为 $O(S \times N + S/T_d \times N \log N)$; 在智能干预层, PI-MADDPG 算法在执行阶段仅执行 actor 网络的前向推理, 其计算复杂度为 $O(N \times L \times d_{\text{hidden}}^2)$, 其中, L 为网络层数, d_{hidden} 为神经网络隐藏层维度, 二者均为常数. 因此, 化简所有常数项后, 框架总的计算复杂度为 $O(N + N \log N)$, 复杂度理论分析表明该框架能够适应大规模人群场景下的实时计算需求.

为验证所提框架在计算实验下的实时性能, 本文以平行执行案例实验的疏散场景为例, 对不同人群规模下仿真与引导干预过程的计算时间开销进行统计与定量分析, 结果如图 18 所示. 设仿真场景中经过的虚拟时间为 T_{sim} (模拟时间), 程序实际运行所消耗的物理时间为 T_{wall} (墙上时间), 定义实时率 (real-time factor, RTF) 为:

$$RTF = \frac{T_{\text{wall}}}{T_{\text{sim}}} \quad (23)$$

RTF 是衡量仿真系统实时性能的指标: RTF 越小计算效率越高; 当 $RTF \leq 1$ 时, 表明系统能够满足实时运行的要求.

实验结果显示, 在不同人群规模条件下, 本文

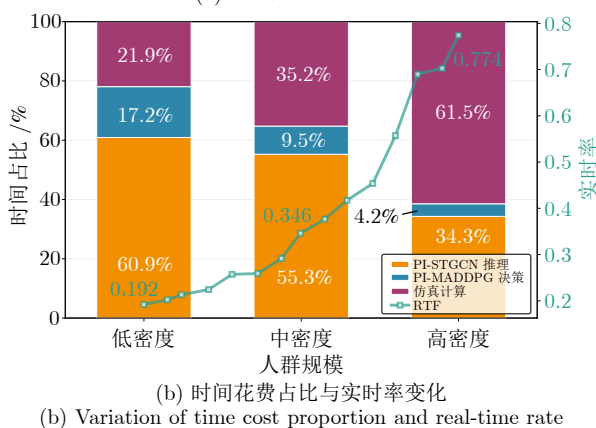
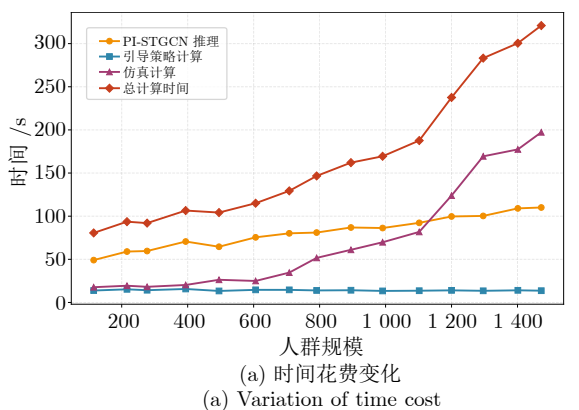


图 18 不同人群规模下的仿真与引导干预计算时间花销
Fig.18 Computation time cost of simulation and guidance intervention under different crowd scales

框架整体均保持了较低的实时率. 具体而言, 随着人群规模逐渐增加, 总计算时间和 RTF 逐渐增长, 增长趋势与理论分析所得的 $O(N + N \log N)$ 复杂度基本一致. 从各部分耗时构成来看, PI-STGCN 推理与仿真计算是主要计算开销来源. 而引导策略计算仅涉及 actor 网络前向推理, 耗时相对稳定且占比较低, 平均推理时间为 14.12 s. 相比之下, 仿真计算时间开销随人群规模的增长最为显著, 其计算时间占比由低密度条件下的 21.9% 升至高密度条件下的 61.5%. 尽管如此, 即使在高密度人群场景下, 框架运行实时率仍能稳定维持在 1 以下, 说明仿真速度不低于实际系统演化的速度, 仍满足实际部署中实时仿真与干预性能的要求.

综上所述, 本文所提出的物理信息机器学习驱动的建模仿真与干预决策框架, 在计算效率层面具备支持大规模人群实时仿真与引导决策的能力. 未来研究可进一步结合并行计算或分布式仿真架构, 增强系统对超大规模人群场景的实时响应与计算扩展能力, 应对更加复杂的社会安全与公共管理需求.

4 结束语

针对微观人群移动管控问题, 本文提出一种物理信息机器学习驱动的建模仿真与干预决策框架. 该框架集成数据感知、物理信息融合建模、动态仿真推演以及智能干预各层级模块, 并基于“平行智能”思想构建人群管理的虚实闭环架构. 在框架中, 本文设计了基于 PI-STGCN 的人群移动导航势能场模型, 并设计 PI-MADDPG 强化学习算法自适应生成引导策略, 有效克服了传统方法中仿真精度不足及干预策略过度依赖人工经验等局限, 为大规模人群管控提供科学计算支撑. 真实人群移动数据集上的实验结果表明, 框架在大规模人群仿真与引导疏散任务的应用中表现出较高的准确性与疏散效率. 通过智能干预层平行执行引导策略, 人群疏散效率显著提升, 相比传统各类引导方法, 疏散时间缩短了 10.2%. 该框架构建从人群移动建模仿真, 到干预策略生成、执行及反馈修正的完整闭环链路, 可在复杂环境中动态调整策略以高效引导人群移动, 对城市公共安全与应急疏散管理具有重要应用价值.

未来研究将进一步优化模型的适应性, 提升其在多样化城市环境中的泛化能力. 在数据驱动层面, 我们将在现有物理交互机制基础上引入社会关系嵌入与群体感知注意力模块, 以增强对心理与社会因素的建模能力; 在机理建模层面, 将通过融合多源数据 (如行人心理特征、社会行为模式、环境天气信息等) 进一步强化仿真真实性与决策干预有效性.

此外,探索复杂群体移动场景中的多智能体协作与协同决策机制,也将是重点推进的研究方向。

附录 A PI-MADDPG 模型训练

在训练阶段,所有引导智能体共享一个经验回放池。算法利用池中存储的经验数据训练 actor 网络和 critic 网络。在每个时间步,每个引导智能体的 actor 网络根据当前状态 $S = \{s_1, s_2, \dots, s_N\}$ 独立生成动作,形成动作集合 $A = \{a_1, a_2, \dots, a_N\}$ 。动作执行后,环境转移至下一时刻状态,并产生即时奖励 $R = \{r_1, r_2, \dots, r_N\}$ 。随后, critic 网络输入当前状态、动作,以及下一时刻的状态,拟合动作价值函数 $Q_i = (s_i, a_i)$ 。完成一次状态-动作-状态循环后,智能体 i 的经验数据以四元组 $[s_i, a_i, s'_i, r_i]$ 的形式存储于经验回放池。每个时间步向池中添加 N 条经验。当池容量饱和时覆盖最早记录。每次训练更新网络时,从经验回放池中随机抽取 K 条经验数据。设 actor 网络为 $f_{actor}(\cdot)$, critic 网络为 $f_{critic}(\cdot)$, $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_N)$ 表示 actor 网络的参数, $\eta = (\eta_1, \eta_2, \dots, \eta_N)$ 表示 critic 网络的参数。则 actor 网络和 critic 网络的输入输出关系为:

$$a_i = f_{actor}(s_i | \varphi_i) \quad (A1)$$

$$Q_i = f_{critic}(s_i, a_i | \eta_i) \quad (A2)$$

$$Q'_i = f_{critic}(s'_i, a'_i | \eta_i) \quad (A3)$$

Q'_i 评估在下一状态 s'_i 下,执行由目标 actor 网络预测的动作 a'_i 的价值。当前动作的目标价值函数为 $r_i + \rho Q'_i$,则 critic 网络当前 Q 值 Q_i 的损失函数如下:

$$L_{critic}(\eta_i) = \frac{1}{K} \sum_{i=1}^K (r_i + \rho Q'_i - Q_i) \quad (A4)$$

actor 网络训练的目标是最大化 critic 网络输出动作的目标 Q 值,智能体 i 的策略梯度计算公式:

$$\nabla_{\varphi_i} J(\varphi_i) = \frac{1}{K} [\nabla_{\varphi_i} \log(f_{actor}(s_i)) \times Q_i(s_i, a_i)] \quad (A5)$$

actor 网络和 critic 网络的参数更新公式如下:

$$\varphi_i = \varphi_i + \xi_A \times \nabla_{\varphi_i} J(\varphi_i) \quad (A6)$$

$$\eta_i = \eta_i - \xi_C \times \nabla_{\eta_i} L_{critic}(\eta_i) \quad (A7)$$

其中, ξ_A 和 ξ_C 分别为 actor 网络和 critic 网络的学习率,目标 actor 和目标 critic 网络的参数 φ' 和 η' 采用软更新的方式更新,设 τ 为软更新参数。则

$$\varphi' = \tau \varphi + (1 - \tau) \varphi' \quad (A8)$$

$$\eta' = \tau \eta + (1 - \tau) \eta' \quad (A9)$$

附录 B 实验参数设置

在仿真实验中,采用留一交叉验证策略:针对 5 个场景设计 5 组实验,每组实验使用其中一个场景的数据进行测试,其余四个场景的数据用于训练。通过多次预实验,综合考虑网络训练时间与稳定性,PI-STGCN 中的时间自注意力机制最终采用具有 4 个头的多头注意力形式实现。时空图卷积部分使用单层 GCN 和单层 TCN 进行时空特征提取,并采用 5 层 MLP 解码器预测行人的初始位移和速度。PINN 由一个 4 层深度卷积网络构成,每个隐藏层包含 64 个神经元,其拟合的运动方程最高阶数设为 4 阶。如图 B1 所示,损失函数中的超参数根据预测移动位置分布的平均位移误差 (average displacement error, ADE) 的数值,通过预实验的网格搜索确定。其中, ω_p 值选择 0.10; ω_c 为动态超参数,初始值设为 0.01,并随训练轮次递增,上限设置为 5。训练批次大小设置为 128,使用随机梯度下降算法训练 200 个轮次,学习率为 0.001。

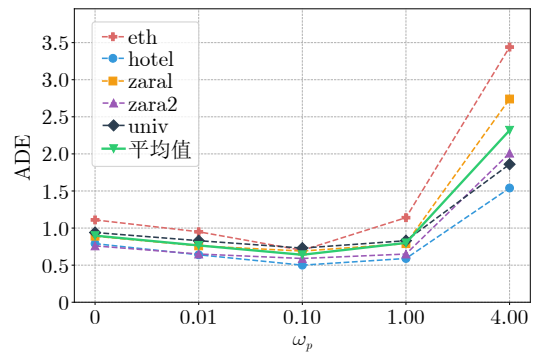


图 B1 不同 ω_p 取值下 PI-STGCN 的 ADE 性能
Fig. B1 ADE performance of PI-STGCN with different ω_p values

仿真模型主要参数设置如表 B1 所示。与 PI-STGCN 输入数据序列的时间间隔保持一致,仿真

表 B1 人群移动仿真实验参数
Table B1 Crowd movement simulation experiment parameters

参数名称	值
$P(t)$ 阶次	4
ω_p	0.10
ω_c	[0.01, 5]
k_{g_i}	10
λ_o	25
λ_h	4
Δt	0.4 s
T_d	12

步长取 0.4 s; 栅格边长为行人站立时通常所占空间的直径长度 (0.4 m). 模型利用 8 个历史步长的数据进行预测, 最大可计算未来 12 个步长的导航势能场, 即 $T_d = 12$. 鉴于数据集中行人间最小间距小于栅格边长, 障碍场与行人交互场的范围边界大小均设置为 1 个栅格单元. 对于数据集中每个场景, 每组参数独立重复运行 10 次仿真实验, 结果取平均值.

在引导干预实验中, PI-MADDPG 的模型参数设置如表 B2 所示. 其中, 环境状态更新时间步长与仿真模型推进步长保持一致. 奖励函数参数通过预实验, 并依据多轮干预结果的反馈及人群移动仿真模型参数的调整确定.

表 B2 人群引导干预实验参数
Table B2 Crowd guidance intervention experiment parameters

参数名称	值
l 阶次	6
$(\beta_{obs}, \beta_{ped})$	(5, 1.2)
w_p	15
w_c	-5
w_d	5
w_t	-2
w_s	25
ξ_A	1×10^{-4}
ξ_C	1×10^{-3}
τ	0.01

在平行执行案例实验中, 本文将 PI-MADDPG 与基于最短路径规划、用户草图以及未引入数据驱动预测的传统 MADDPG 等基线方法进行了对比分析, 各类方法的描述与核心特点如下:

1) 基于最短路径规划的方法依据环境拓扑结构为引导智能体规划静态最短路径, 其特点是路径固定、静态干预, 未考虑行人之间的动态交互行为.

2) 基于用户草图的方法通过人工标注宏观引导方向或区域, 使智能体按预设路径移动. 该方法高度依赖管理者的先验知识, 响应具有滞后性, 难以适应复杂动态场景的实际需求.

3) 传统 MADDPG 方法采用多智能体强化学习框架进行决策, 但未引入物理信息机器学习驱动的移动预测机制. 该方法虽具备动态干预能力, 但由于缺乏对人群动态的显式建模, 引导精准度有限.

参考文献

1 Yue Shi-Yao. Analysis on stampede accidents of large-scale spon-

aneous congregation. *City and Disaster Reduction*, 2023(3): 17-21

(岳诗瑶. 大型自发聚集踩踏事故探析——从梨泰院踩踏事故谈起. *城市与减灾*, 2023(3): 17-21)

2 Sun Xu-Bin, Dong Hai-Rong, Ning Bin, Gao Tong-Xin, Kong Qing-Jie. ACP-based emergency evacuation system. *Acta Automatica Sinica*, 2014, **40**(1): 16-23

(孙绪彬, 董海荣, 宁滨, 高童欣, 孔庆杰. 基于 ACP 方法的应急疏散系统研究. *自动化学报*, 2014, **40**(1): 16-23)

3 Zhang J B, Zhu J, Dang P, Wu J L, Zhou Y X, Li W L, et al. An improved social force model (ISFM)-based crowd evacuation simulation method in virtual reality with a subway fire as a case study. *International Journal of Digital Earth*, 2023, **16**(1): 1186-1204

4 Dang H T, Gaudou B, Verstaevl N. A literature review of dense crowd simulation. *Simulation Modelling Practice and Theory*, 2024, **134**: Article No. 102955

5 Echeverri L C, Auberlet J M, Hubert J P. Modeling collision avoidance behavior with zero-speed pedestrians. *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**(8): 9608-9617

6 Subaih R, Tordeux A. Modeling pedestrian single-file movement: Extending the interaction to the follower. *Physica A: Statistical Mechanics and Its Applications*, 2024, **633**: Article No. 129394

7 Blue V J, Embrecht M J, Adler J L. Cellular automata modeling of pedestrian movements. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Orlando, USA: IEEE, 1997. 2320-2323

8 Bazior G, Wąs J, Pałka D. Pedestrian dynamics model for high densities. *Expert Systems With Applications*, 2025, **272**: Article No. 126775

9 Helbing D, Molnár P. Social force model for pedestrian dynamics. *Physical Review E*, 1995, **51**(5): 4282-4286

10 Zhao R Y, Wei B Y, Han C F, Jia P, Zhu W J, Li C L, et al. Improved crowd dynamics analysis considering physical contact force and panic emotional propagation. *IEEE Transactions on Intelligent Transportation Systems*, 2025, **26**(2): 1840-1851

11 Cheney S. Flow tiles. In: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Grenoble, France: ACM, 2004. 233-242

12 Zhu L J, Xu X M, Wang J, Chen J H, Ma Z J, Wang Q, et al. An enhanced multi-attribute crowd evacuation model incorporating the effects of environmental impact factors. *Heliyon*, 2024, **10**(4): Article No. e26299

13 Du Quan-Cheng, Wang Xiao, Li Ling-Xi, Ning Huan-Sheng. Key problems and progress of pedestrian trajectory prediction methods: The state of the art and prospects. *Chinese Journal of Intelligent Science and Technology*, 2023, **5**(2): 143-162 (杜泉成, 王晓, 李灵犀, 宁焕生. 行人轨迹预测方法关键问题研究: 现状及展望. *智能科学与技术学报*, 2023, **5**(2): 143-162)

14 Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L, Savarese S. Social LSTM: Human trajectory prediction in crowded spaces. In: Proceedings of the 39th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 961-971

15 Gupta A, Johnson J, Fei-Fei L, Savarese S, Alahi A. Social GAN: Socially acceptable trajectories with generative adversarial networks. In: Proceedings of the 41st IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 2255-2264

16 Yan D P, Ding G Y, Huang K X, Huang T Y. Generating natural pedestrian crowds by learning real crowd trajectories through a Transformer-based GAN. *The Visual Computer*, 2025, **41**(2): 1079-1096

17 Mohamed A, Qian K, Elhoseiny M, Claudel C. Social-STGCNN: A social spatio-temporal graph convolutional neural network for human trajectory prediction. In: Proceedings of the 43rd IEEE Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE, 2020. 14424-14432

- 18 Huang Y F, Bi H K, Li Z X, Mao T L, Wang Z Q. STGAT: Modeling spatial-temporal interactions for human trajectory prediction. In: Proceedings of the 17th IEEE International Conference on Computer Vision. Seoul, Korea: IEEE, 2019. 6272–6281
- 19 Chen W X, Sang H F, Wang J Y, Zhao Z S. WTGCN: Wavelet transform graph convolution network for pedestrian trajectory prediction. *International Journal of Machine Learning and Cybernetics*, 2024, **15**(12): 5531–5548
- 20 Yu C J, Ma X, Ren J W, Zhao H Y, Yi S. Spatio-temporal graph Transformer networks for pedestrian trajectory prediction. In: Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer, 2020. 507–523
- 21 Shi L S, Wang L, Zhou S P, Hua G. Trajectory unified Transformer for pedestrian trajectory prediction. In: Proceedings of the 19th IEEE International Conference on Computer Vision. Paris, France: IEEE, 2023. 9675–9684
- 22 Gu T P, Chen G Y, Li J L, Lin C Z, Rao Y M, Zhou J, et al. Stochastic trajectory prediction via motion indeterminacy diffusion. In: Proceedings of the 45th IEEE Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE, 2022. 17113–17122
- 23 Wang K L, Tsao L W, Wu J C, Shuai H H, Cheng W H. Traj-Fine: Predicted trajectory refinement for pedestrian trajectory forecasting. In: Proceedings of the 47th IEEE Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA: IEEE, 2024. 4483–4492
- 24 Zhang C Y, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, 2021, **64**(3): 107–115
- 25 Zhang G Z, Yu Z H, Jin D P, Li Y. Physics-infused machine learning for crowd simulation. In: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA: ACM, 2022. 2439–2449
- 26 Ni Ling-Jia, Huang Xiao-Xia, Li Hong-Ga, Zhang Zi-Bo. Research on fire emergency evacuation simulation based on cooperative deep reinforcement learning. *Journal of System Simulation*, 2022, **34**(6): 1353–1366
(倪凌佳, 黄晓霞, 李红杏, 张子博. 基于协作式深度强化学习的火灾应急疏散仿真研究. 系统仿真学报, 2022, **34**(6): 1353–1366)
- 27 Zhang K, Zhu D D, Xu Q H, Zhou H, Peng X M. APF-DQN: Adaptive objective pathfinding via improved deep reinforcement learning among building fire hazard. In: Proceedings of the 33rd International Conference on Artificial Neural Networks Artificial Neural Networks and Machine Learning. Lugano, Switzerland: Springer, 2024. 265–279
- 28 Li X J, Liu H, Li J Q, Li Y. Deep deterministic policy gradient algorithm for crowd-evacuation path planning. *Computers & Industrial Engineering*, 2021, **161**: Article No. 107621
- 29 Wang H Y, Liu H, Li W H. Crowd evacuation path planning and simulation method based on deep reinforcement learning and repulsive force field. *Applied Intelligence*, 2025, **55**(4): Article No. 297
- 30 Yang Jing, Wang Xiao, Wang Yu-Tong, Liu Zhong-Min, Li Xiao-Shuang, Wang Fei-Yue. Parallel intelligence and CPSS in 30 years: An ACP approach. *Acta Automatica Sinica*, 2023, **49**(3): 614–634
(杨静, 王晓, 王雨桐, 刘忠民, 李小双, 王飞跃. 平行智能与 CPSS: 三十年发展的回顾与展望. 自动化学报, 2023, **49**(3): 614–634)
- 31 Wang Fei-Yue. Computational theory and method on complex system. *China Basic Science*, 2004, **6**(5): 3–10
(王飞跃. 关于复杂系统研究的计算理论与方法. 中国基础科学, 2004, **6**(5): 3–10)
- 32 Zhao Y, Zhu Z Q, Chen B, Qiu S H, Huang J C, Lu X, et al. Toward parallel intelligence: An interdisciplinary solution for complex systems. *The Innovation*, 2023, **4**(6): Article No. 100521
- 33 Korbmacher R, Tordeux A. Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Transportation Systems*, 2022, **23**(12): 24126–24144
- 34 Colas A, van Toll W, Zibrek K, Hoyet L, Olivier A H, Pettré J. Interaction fields: Intuitive sketch-based steering behaviors for crowd simulation. *Computer Graphics Forum*, 2022, **41**(2): 521–534
- 35 Macal C M, North M J. Agent-based modeling and simulation. In: Proceedings of the Winter Simulation Conference. Austin, USA: IEEE, 2009. 86–98
- 36 Hewitt C, Inman J. DAI betwixt and between: From ‘intelligent agents’ to open systems science. *IEEE Transactions on Systems, Man, and Cybernetics*, 1991, **21**(6): 1409–1419
- 37 Karniadakis G E, Kevrekidis I G, Lu L, Perdikaris P, Wang S F, Yang L. Physics-informed machine learning. *Nature Reviews Physics*, 2021, **3**(6): 422–440
- 38 Zhang D K, Guo L, Karniadakis G E. Learning in modal space: Solving time-dependent stochastic PDEs using physics-informed neural networks. *SIAM Journal on Scientific Computing*, 2020, **42**(2): A639–A665
- 39 Liu Yan-Fei, Li Chao, Wang Zhong, Wang Jie-Ling. Research progress on multi-agent deep reinforcement learning and scalability. *Computer Engineering and Applications*, 2025, **61**(4): 1–24
(刘延飞, 李超, 王忠, 王杰铃. 多智能体深度强化学习及可扩展性研究进展. 计算机工程与应用, 2025, **61**(4): 1–24)
- 40 Luo Biao, Hu Tian-Meng, Zhou Yu-Hao, Huang Ting-Wen, Yang Chun-Hua, Gui Wei-Hua. Survey on multi-agent reinforcement learning for control and decision-making. *Acta Automatica Sinica*, 2025, **51**(3): 510–539
(罗彪, 胡天萌, 周育豪, 黄廷文, 阳春华, 桂卫华. 多智能体强化学习控制与决策研究综述. 自动化学报, 2025, **51**(3): 510–539)
- 41 Lowe R, Wu Y, Tamar A, Harb J, Abbeel P, Mordatch I. Multi-agent actor-critic for mixed cooperative-competitive environments. In: Proceedings of the 31st International Conference on Neural Information Processing System. Long Beach, USA: Curran Associates Inc., 2017. 6382–6393
- 42 Li Yi-Chun, Liu Ze-Jiao, Hong Yi-Tian, Wang Ji-Chao, Wang Jian-Rui, Li Yi, et al. Multi-agent reinforcement learning based game: A survey. *Acta Automatica Sinica*, 2025, **51**(3): 540–558
(李艺春, 刘泽娇, 洪艺天, 王继超, 王健瑞, 李毅, 等. 基于多智能体强化学习的博弈综述. 自动化学报, 2025, **51**(3): 540–558)
- 43 Jiang B, Zhang Z Y, Lin D D, Tang J, Luo B. Semi-supervised learning with graph learning-convolutional networks. In: Proceedings of the 42nd IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 11313–11320
- 44 Pellegrini S, Ess A, Schindler K, van Gool L. You’ll never walk alone: Modeling social behavior for multi-target tracking. In: Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 261–268
- 45 Lerner A, Chrysanthou Y, Lischinski D. Crowds by example. *Computer Graphics Forum*, 2007, **26**(3): 655–664
- 46 Bönsch A, Barton S J, Ehret J, Kuhlen T W. Immersive sketching to author crowd movements in real-time. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents. Scotland, UK: ACM, 2020. Article No. 11



郭润康 国防科技大学系统工程学院博士研究生。2024 年获得国防科技大学硕士学位。主要研究方向为复杂系统建模与仿真。

E-mail: guorunkangnudt@mudt.edu.cn
(GUO Run-Kang Ph.D. candidate at the College of Systems Engineering, National University of Defense Technology. He received his master degree from National University of Defense Technology in 2024. His research interests include complex system modeling and simulation.)



朱正秋 国防科技大学数智建模与仿真国家级重点实验室副研究员. 主要研究方向为复杂系统建模与仿真, 群智计算, 具身智能. 本文通信作者.

E-mail: zhuzhengqiu12@nudt.edu.cn
(**ZHU Zheng-Qiu** Associate professor at the State Key Laboratory

of Digital-Intelligent Modeling and Simulation, National University of Defense Technology. His research interests include complex system modeling and simulation, crowdsensing computing, and embodied intelligence. Corresponding author of this paper.)



秦 龙 国防科技大学数智建模与仿真国家级重点实验室副研究员. 2014 年获得国防科技大学博士学位. 主要研究方向为复杂系统建模与仿真.

E-mail: qldbx2007@sina.com

(**QIN Long** Associate professor at the State Key Laboratory of Digital-

Intelligent Modeling and Simulation, National University of Defense Technology. He received his Ph.D. degree from National University of Defense Technology in 2014. His research interests include complex system modeling and simulation.)



艾 川 国防科技大学数智建模与仿真国家级重点实验室讲师. 主要研究方向为高性能仿真, 社会仿真.

E-mail: aichuan@nudt.edu.cn

(**AI Chuan** Lecturer at the State Key Laboratory of Digital-Intelligent Modeling and Simulation, National

University of Defense Technology. His research interests include high performance simulation and social simulation.)



尹全军 国防科技大学数智建模与仿真国家级重点实验室研究员. 2005 年获得国防科技大学博士学位. 主要研究方向为行为建模, 云仿真.

E-mail: yin_qaunjun@163.com

(**YIN Quan-Jun** Professor at the State Key Laboratory of Digital-Intelligent

Modeling and Simulation, National University of Defense Technology. He received his Ph.D. degree from National University of Defense Technology in 2005. His research interests include behavior modeling and cloud simulation.)



叶佩军 中国科学院自动化研究所副研究员. 2013 年获得中国科学院大学博士学位. 主要研究方向为多智能体系统, 复杂系统建模与控制, 智能交通. E-mail: peijun.ye@ia.ac.cn

(**YE Pei-Jun** Associate professor at the Institute of Automation, Chinese

Academy of Sciences. He received his Ph.D. degree from University of Chinese Academy of Sciences in 2013. His research interests include multi-agent systems, modeling and control of complex systems, and intelligent transportation.)



王飞跃 中国科学院自动化研究所研究员. 主要研究方向为智能系统和复杂系统的建模、分析与控制.

E-mail: feiyue.wang@ia.ac.cn

(**WANG Fei-Yue** Professor at the Institute of Automation, Chinese Academy of Sciences. His research

interests include modeling, analysis, and control of intelligent systems and complex systems.)