

基于多智能体强化学习的分布式电源集群协同控制与出力优化

韩建刚

河南袍尔电气有限公司, 河南省许昌市, 461001;

摘要: 随着能源结构的转型和智能化技术的发展, 分布式电源系统在现代电力网络中的重要性日益凸显。然而, 分布式电源的间歇性和波动性给电网的稳定运行带来了巨大挑战。传统的集中式控制方法在面对大规模分布式电源集群时, 往往表现出计算复杂度高、通信负担重以及扩展性差等问题。因此, 如何设计一种高效且灵活的协同控制策略成为亟待解决的关键问题。多智能体强化学习作为一种新兴的人工智能技术, 为分布式电源集群的协同优化提供了全新的思路。通过多个智能体之间的交互与学习, 可以在动态环境中实现高效的决策和优化, 从而提升系统的整体性能。这一研究方向不仅具有理论创新价值, 还具备广泛的实际应用前景。

关键词: 多智能体强化学习; 分布式电源集群; 协同控制; 出力优化

DOI: 10. 64216/3080-1508. 25. 08. 038

引言

随着全球能源需求的不断增长和环境保护意识的提升, 构建清洁、高效、可持续的能源系统已成为各国的重要战略目标。分布式电源因其灵活部署、绿色环保等优势, 在能源体系中扮演着越来越重要的角色。然而, 分布式电源的分散性、不确定性和多样性也对电力系统的控制与优化提出了更高的要求。传统的集中式管理方式难以满足复杂多变的实际需求, 尤其是在面对大规模分布式电源集群时, 其局限性愈发明显。因此, 探索一种能够适应动态环境、具备良好扩展性和鲁棒性的新型控制方法显得尤为必要。多智能体强化学习技术的引入, 为解决这些问题提供了新的可能性, 其通过分布式决策和协同优化的方式, 能够在复杂的电力网络中实现更高效的资源利用和更稳定的系统运行。

1 分布式电源集群概述

分布式电源集群并非简单的 DG 集合, 而是通过通信网络、控制单元与储能系统, 将光伏、风电、微型燃气轮机、储能装置等异构单元聚合, 形成具备“自主协调、灵活并网、按需响应”能力的电力子系统。其核心结构包含三层: 底层为 DG 单元与负荷节点, 是能量生产与消耗的基础; 中层为区域控制单元, 负责局部数据采集与实时调节; 顶层为集群协调中心, 统筹全局优化目标。这种分层结构既避免了单一单元的随机性干扰, 又通过聚合效应提升了集群与主网的交互能力——例如, 当某一光伏单元因云层遮挡出力骤降时, 储能系统可快速放电补能, 邻近燃气轮机可微调出力, 通过中层

控制单元的协同, 避免对主网造成冲击。

DG 集群的关键特征决定了其控制需求的特殊性: 一是异构性, 不同 DG 的出力特性差异显著 (光伏日间波动、风电间歇性、燃气轮机可控性), 需通过协同机制实现能源互补; 二是动态性, 负荷需求的时段变化、气象条件的实时波动, 要求集群具备快速响应能力; 三是分布式特性, 集群内单元地理分散, 传统集中式控制的通信延迟会降低调节精度; 四是并网交互性, 集群作为整体与主网进行功率交换, 需满足主网的电压、频率约束, 同时最大化自身新能源消纳率。从电力系统价值看, DG 集群不仅能实现“削峰填谷”——用电高峰时释放储能与可控 DG 出力, 低谷时吸收过剩新能源充电, 还能提升供电可靠性, 当主网故障时, 具备孤岛运行能力的集群可保障关键负荷供电, 是构建新型电力系统的重要支撑。

2 多智能体强化学习在分布式电源集群中的问题

2.1 智能体通信耦合与目标冲突问题

多智能体系统的协同依赖高效信息交互, 但 DG 集群的通信环境存在天然局限: 一是通信延迟与丢包, 集群内单元地理分散, 无线通信易受天气、地形干扰, 有线通信成本高且扩展性差, 导致智能体间状态信息传递滞后, 例如某 DG 智能体已完成出力调整, 但其状态反馈未及时传递给邻近智能体, 可能引发重复调节或调节不足; 二是信息不对称, 各智能体的感知范围有限, 仅

能获取局部运行数据（如自身出力、周边负荷），难以掌握全局状态（如主网功率需求、远端 DG 出力），易导致局部决策与全局目标脱节；三是目标冲突，各 DG 单元可能存在局部利益诉求（如业主追求自身发电收益最大化），与集群全局目标（如降低弃风弃光率、维持电网稳定）产生冲突——例如，某风电智能体为获取更多补贴，可能在电网负荷低谷时仍维持高出力，导致集群弃风率上升，甚至引发主网电压越限。这种通信耦合与目标冲突，会削弱 MARL 的协同效果，导致策略优化陷入局部最优。

2.2 动态环境适应性与策略滞后问题

DG 集群运行环境动态性远超传统电力系统，核心扰动源有三类：一是新能源出力波动，光伏受辐照度、温度影响，风电受风速、风向影响，短时间出力预测误差可能超 10%；二是负荷需求突变，工业负荷启停、居民负荷时段性变化使集群功率平衡需求快速调整；三是电网拓扑变化，主网检修、线路故障等改变集群并网接口条件，如主网线路断开，集群需调整出力适应新功率传输约束。

MARL 策略更新依赖“试错-学习”循环，但传统 MARL 采用静态奖励函数与固定更新频率，难以适应动态环境。一方面，静态奖励函数无法实时匹配环境变化，如用电高峰和低谷应优先奖励不同策略，固定奖励函数会使策略方向与需求脱节；另一方面，策略更新有滞后性，环境突变时智能体需多轮交互调整策略，而电力系统对响应速度要求高，滞后策略可能致系统失稳。此外，部分智能体故障会使系统拓扑与能力变化，传统 MARL 缺乏故障场景鲁棒性训练，易出现“策略失效”问题。

2.3 多优化目标的维度平衡与权重分配问题

DG 集群的控制与出力优化涉及多维度目标且相互冲突：经济性目标要求降低发电成本和弃能损失；安全性目标要求维持集群电压、频率稳定，避免线路过流；环保性目标要求减少碳排放。这些目标的冲突给 MARL 带来挑战，单一目标优化会致系统失衡，如过度追求经济性会引发电压波动，过度强调安全性会降低新能源消纳率。

传统 MARL 多采用加权求和法整合多目标，存在两大问题：一是权重设定主观化，依赖专家经验，难以适应不同运行场景；二是权重静态化，无法实时响应环境变化，如主网碳排放政策调整时无法自动适配。此外，

部分目标难以量化，无法纳入奖励函数，导致 MARL 优化方向偏差，难以实现全局平衡。

3 基于多智能体强化学习的分布式电源集群协同控制与出力优化策略

3.1 分层式多智能体架构设计策略

针对智能体通信耦合与目标冲突问题，构建“全局协调-区域控制-局部执行”三层多智能体架构，实现“集中决策与分布式执行”结合。顶层全局协调智能体（GCA）部署于集群协调中心，负责全局目标优化，采集主网调度指令与集群运行状态，制定全局优化目标并拆解为各区域子目标。中层区域控制智能体（RCA）按地理或能源类型划分区域，接收 GCA 子目标，结合区域内 DG、储能运行状态制定区域协同策略，同时反馈区域执行状态。底层局部执行智能体（LEA）对应单个 DG、储能或负荷单元，接收 RCA 策略指令实现实时出力调节，并反馈局部状态。

该架构有三大核心优势：一是降低通信负担，GCA 与 RCA 间、RCA 与 LEA 间传递特定指令，避免全局数据传输；二是化解目标冲突，GCA 拆解全局目标为区域任务，RCA 协调局部利益与区域目标，如调整储能策略保障区域弃能率达标；三是提升容错性，某一层级智能体故障时，其他层级可临时接管部分功能，如 GCA 向故障区域 LEA 下发简化指令维持基本运行。

3.2 基于联邦强化学习的策略更新机制

为解决动态环境适应性与数据隐私问题，引入联邦强化学习（FedRL）构建策略更新机制，实现“本地训练-全局聚合-分布式执行”闭环优化。核心逻辑为：各 LEA 作为联邦学习客户端，基于本地运行数据进行强化学习训练，更新局部策略模型；RCA 作为区域联邦服务器，收集 LEA 局部模型参数，用联邦平均算法（FedAvg）聚合得区域全局模型并下发，实现区域策略协同；GCA 作为集群联邦服务器，收集 RCA 区域全局模型参数，聚合得集群全局模型下发，保障集群策略一致。

该机制创新点有：一是保护数据隐私，LEA 仅上传模型参数，不泄露原始发电数据，符合电力系统数据安全要求；二是降低通信成本，模型参数数据量小，缓解集群通信压力，适应动态环境快速更新需求；三是提升泛化能力，多客户端数据聚合训练使集群全局模型覆盖不同特征，避免“过拟合”，如光伏 LEA 局部模型经联邦聚合提升对气象波动适应性。此外，采用“增量更新”

策略,环境变化小时,仅 LEA 更新局部模型,RCA 无需重聚,缩短更新耗时,适应电力系统实时性要求。

3.3 动态多目标奖励函数设计策略

针对多目标平衡问题,构建基于“场景识别-权重自适应-多维度量化”的动态奖励函数以实现目标权重实时优化。首先,引入场景识别模块,用深度学习模型(如 LSTM)分析集群运行状态(负荷水平、新能源出力波动、主网约束),自动划分运行场景,如“高峰-高波动场景”(用电高峰、风电出力波动 $>15\%$)、“低谷-低波动场景”(用电低谷、光伏出力波动 $<5\%$)、“故障场景”(主网线路故障、DG 停机)。其次,设计自适应权重算法,基于场景特征动态调整目标权重:“高峰-高波动场景”提升安全性权重至 40%,降低经济性和环保性权重;“低谷-低波动场景”提升经济性和环保性权重,降低安全性权重;“故障场景”提升安全性权重至 60%,降低其他目标权重。

奖励函数量化设计覆盖多维度目标:安全性奖励用“偏差惩罚制”,如电压偏差不同范围对应不同奖励;经济性奖励用“收益-成本制”,如弃能率、燃气轮机耗气量达标有奖励;环保性奖励用“碳减排激励制”,如新能源出力占比达标有奖励。同时,引入“协同奖励项”,策略调整促进区域协同额外奖励+5,鼓励智能体主动协同。这种动态奖励函数能实时匹配系统需求,避免单一目标优化弊端,引导 MARL 向全局最优收敛。

3.4 基于状态预测的超前协同控制策略

为解决动态环境响应滞后问题,融合“短期预测-超前决策-实时修正”构建超前协同控制策略,提升系统前瞻性。首先,构建多源状态预测模型,采用“物理模型+数据驱动模型”融合方法:用注意力机制-LSTM 模型结合物理参数与历史数据预测光伏未来 15-30 分钟出力曲线,误差控制在 8%以内;用 XGBoost 模型结合用户习惯等预测负荷未来 15-30 分钟变化;通过 GCA 接收主网短期指令明确集群并网边界条件。

接着,将预测结果作为 MARL 的“超前状态输入”,各智能体基于未来状态提前调整策略,如预测光伏出力下降,储能 LEA 提前准备放电、调整阈值,燃气轮机 LEA 提前提升备用出力,RCA 提前协调制定协同补能方案。同时,引入“实时修正机制”,每 5 分钟采集实际与预测偏差,若偏差 $>10\%$ 则触发策略重优化,调整各 LEA 出力指令。

这种超前控制策略将 MARL 的“被动响应”变为“主动预判”,缩短响应时间,减少动态环境波动冲击,如风电出力骤降前超前调整储能放电,可将电压波动幅度从 $\pm 4\%$ 降至 $\pm 2\%$,提升系统稳定性。

3.5 故障自适应协同容错策略

针对智能体故障致策略失效问题,构建“故障检测-隔离-任务重构”容错机制保障集群系统鲁棒性。首先,设计分布式故障检测模块,各智能体通过“状态交互+异常识别”诊断故障:LEA 实时向 RCA 发送运行状态,RCA 对比相邻 LEA 状态反馈,结合异常检测算法判断智能体是否故障,检测耗时控制在 10 秒以内。其次,执行故障隔离操作,RCA 将故障 LEA 标记为“离线状态”,切断其与其他智能体的策略交互,避免故障扩散,如某储能 LEA 故障后,RCA 停止下发指令并通知邻近 LEA 忽略其状态反馈。最后,实施任务重构策略,RCA 基于“能力匹配”原则,将故障 LEA 的控制任务分配给健康智能体,若故障单元为光伏 LEA,将负荷平衡任务分配给邻近储能 LEA;若故障单元为 RCA,由 GCA 临时接管控制任务,同时启动 RCA 备用单元切换。

此外,为提升 MARL 容错能力,在策略训练阶段引入“故障场景注入”,模拟不同类型智能体故障,让智能体学习故障应对策略,形成“故障适应型”策略模型。例如,训练中定期模拟光伏 LEA 故障,引导储能与燃气轮机 LEA 学习协同补能策略,使系统在实际故障时快速响应,保障集群运行的连续性与稳定性。

4 结语

本文为分布式电源集群的协同控制与出力优化提供了新的理论和方法,对于推动分布式电源在智能电网中的应用具有重要意义。随着人工智能技术的不断进步,相信多智能体强化学习将在电力系统等领域发挥更加重要的作用。

参考文献

- [1] 赵永杰. 基于多智能体强化学习的分布式发电系统协同优化[J]. 电力系统自动化, 2023, 44(11): 12-13.
- [2] 李慧敏. 多智能体强化学习在分布式电源调度中的应用研究[J]. 电力科学与技术, 2023, 43(4): 17-19.
- [3] 胡瑞强. 基于多智能体强化学习的分布式电源集群出力优化策略研究[J]. 电力系统保护与控制, 2023, 49(1): 23-25.