

# 基于深度强化学习的主动配电网动态重构综述

江昌旭, 郭辰, 刘晨曦, 林俊杰, 邵振国

(福州大学电气与自动化工程学院(福建省电器智能化工程技术研究中心), 福州 350108)

**摘要:** 随着双碳目标的快速发展, 大量以风电、光伏为代表的分布式电源接入配电网, 这将进一步加剧电源出力的间歇性与波动性。主动配电网动态重构属于一个复杂的高维混合整数非线性随机优化问题, 传统算法在解决该问题的过程中存在着诸多不足之处。而深度强化学习算法结合了深度学习与强化学习的优势, 非常适用于制定当前备受关注的主动配电网动态重构策略。该文首先对新型电力系统主动配电网特征进行总结, 并对当前主动配电网动态重构研究在构建数学模型方面所取得的进展以及所面临的挑战进行了深入分析。其次, 对配电网动态重构编码方式进行了探讨, 并对深度强化学习算法进行了系统地综述。进而, 重点分析了现有算法在处理主动配电网动态重构时的不足之处, 并对深度强化学习算法在主动配电网动态重构方面的研究现状与优势进行了总结与概括。最后, 对主动配电网动态重构的未来研究方向进行了展望。

**关键词:** 主动配电网; 动态重构; 深度强化学习; 编码方式; 机器学习; 人工智能

## Review of Active Distribution Network Dynamic Reconfiguration Based on Deep Reinforcement Learning

JIANG Changxu, GUO Chen, LIU Chenxi, LIN Junjie, SHAO Zhenguo

(College of Electrical Engineering and Automation, Fuzhou University(Fujian Smart Electrical Engineering Technology Research Center), Fuzhou 350108, China)

**Abstract:** With the rapid development of dual-carbon targets, many distributed power sources, represented by wind power and photovoltaics, are being connected to distribution networks. This will further exacerbate the intermittency and volatility of power output. Dynamic reconfiguration of active distribution networks constitutes a complex, high-dimensional, mixed-integer, nonlinear, and stochastic optimization problem. Traditional algorithms exhibit numerous shortcomings in addressing this issue. By integrating the advantages of both deep learning and reinforcement learning, the deep reinforcement learning algorithm is highly suitable for formulating dynamically reconfigurable strategies for active distribution networks, which are currently of great concern. This paper first summarizes the characteristics of the active distribution network of the new generation power system, and analyzes the progress and challenges of the current research on the dynamic reconfiguration of the active distribution network in mathematical models. Secondly, the coding method of the distribution network dynamic reconfiguration is discussed, and the deep reinforcement learning algorithm is systematically reviewed. Furthermore, the shortcomings of the existing algorithms in dealing with the dynamic reconfiguration of the active distribution network are analyzed, and the research status and advantages of the deep reinforcement learning algorithm in the dynamic reconfiguration of the active distribution network are summarized. Finally, the future research directions for the dynamic reconfiguration of active distribution networks are presented.

**Key words:** active distribution network; dynamic reconfiguration; deep reinforcement learning; encoding method; machine learning; artificial intelligence

## 0 引言

分布式电源(distributed generation, DG)能够有

效减少碳排放, 助力实现“碳达峰、碳中和”的双碳目标。因此, 大量风机和光伏发电等 DG 被广泛集成到电网中, 使得传统配电网正在向主动配电网转变<sup>[1]</sup>。当主动配电网中的大规模 DG 出力较大时, 配电网的基础潮流将发生显著变化, 甚至可能导致功率逆向输送与电压越上限, 进而对网损和电压特性产生严重的不利影响, 导致网络性能显著劣化。

配电网重构通过开断/闭合联络开关的方式来优化配电网结构与运行方式,不仅能有效提高清洁能源的消纳比例,还能有效改善配电网运行指标,是一种非常有效的配电网管理和运行技术<sup>[2]</sup>。

目前大部分重构研究仅在某一时间断面进行静态重构,无法及时响应负荷变化。而动态重构能够根据负荷与 DG 出力随时间的变化,实时地调整配电网的开关状态,从而有效地提高配电网的灵活性和响应速度<sup>[3]</sup>。因此,目前主动配电网动态重构(active distribution network dynamic reconfiguration, ADNDR)得到了越来越广泛的关注。然而,ADNDR 模型中通常包含离散变量(如联络线开/断)、连续变量(如电压、电流和潮流)以及随机变量(如 DG 出力、电力负荷),并且其约束条件中存在大量的非线性项(如二次方项和三角函数)。因此,ADNDR 是一个复杂、高维混合整数非线性随机优化问题,并且该问题的规模随着配电网规模的增加而急剧增大<sup>[4]</sup>。

目前,大多数 ADNDR 算法都是基于模型驱动的优化算法<sup>[5-6]</sup>。该类算法需要准确的配电网参数,由于配电网维度的增加和季节性的天气变化,准确的参数通常很难获得。与此同时,这类算法在解决随机不确定性问题时存在求解速度与准确度难以同时兼顾的矛盾。此外,传统基于模型驱动算法的计算量随着网络规模的扩大呈指数级增长,导致该类算法难以适用于大规模配电网的动态重构。为了快速且准确地对 ADNDR 模型进行求解,部分学者采用了基于数据驱动的算法。其中强化学习(reinforcement learning, RL)和深度强化学习(deep reinforcement learning, DRL)是典型的免模型数据驱动算法,它们通过观察环境信息和环境的反馈来学习 ADNDR 策略,实现从状态到动作的映射。相比于传统基于模型驱动的方法,DRL 算法在解决主动配电网动态重构问题的过程中具有以下 3 个方面的优势:(1)DRL 算法无需借助配电网参数进行模型构建,DRL 智能体通过与环境的交互直接从环境中学习,属于一种端对端的免模型算法;(2)DRL 算法具有较强的学习能力和泛化能力,对于具有多重不确定因素的配电网环境具有较强的适应性;(3)DRL 算法能够通过考虑长期回报学习最优的重构策略,具有一定的动态适应性,能够较好地解决主动配电网动态重构这类序贯决策优化问题,从而较好地满足配电网的长期运行需求<sup>[7]</sup>。综上所述,DRL 算法作为一种兼顾求解效率、求解精度、泛化性能与全局

搜索能力的算法,能够有效地实现对不同规模 ADNDR 模型的求解,得到了越来越广泛的关注。

本文首先总结了新型电力系统主动配电网特征,并对现有 ADNDR 研究在构建数学模型方面的进展以及存在的问题进行了分析;然后,对 ADNDR 编码方式进行了总结和分析;其次,对多种不同的 DRL 算法进行综述;接着,系统性地对 ADNDR 策略进行了总结,并分析了传统算法在求解 ADNDR 时的不足之处;进而,对 DRL 算法在 ADNDR 方面已有的研究成果与优势进行了总结与概括;最后,对 ADNDR 未来的研究方向进行了展望。

## 1 主动配电网动态重构数学模型

区别于传统配电网,主动配电网具备下列特征:(1)配置一定比例的分布式可控资源,分布式可控资源使得动态重构策略的制定更加灵活可控,可以实时调整电网的供需平衡;(2)具有较完善的可观可控水平,高可观可控性为动态重构提供了更多实时、准确的信息,有助于制定更加精准的重构策略;(3)具有实现协调优化的管控中心,管控中心可以在动态重构策略制定过程中,综合考虑配电网的多种优化目标与约束条件,并通过优化算法计算出最优的重构方案;(4)具有可灵活调节的拓扑结构,这为重构策略的制定提供了更多选择。

正是基于主动配电网的这些特性,其动态重构领域的研究在构建数学模型方面已取得了显著进展,具体体现在以下几个方面。

### 1) 优化目标及约束条件多元化

由于配电网结构的日益复杂、不同用户的个性化需求以及可持续发展的要求等因素,ADNDR 数学模型所考虑的优化目标及约束条件日益多元化。配电网运行的经济性是 ADNDR 的优化目标之一,其通过网络损耗<sup>[8]</sup>、负荷均衡度(load balancing, LB)<sup>[9]</sup>、开关操作次数<sup>[10]</sup>、综合费用以及系统平均削减功率(system average curtailed power, SACP)<sup>[11]</sup>进行量化。

此外,配电网运行的可靠性也是 ADNDR 的优化目标之一,配电网故障恢复重构主要通过缩短故障发生时间以及降低故障发生频率的方式,来提高配电网运行可靠性。配电网系统发生故障的频率与时间常通过系统平均中断次数指标(system average interruption frequency index, SAIFI)<sup>[12]</sup>、系统平均中断持续时间指标(system average interruption dura-

tion index, SAIDI)<sup>[13]</sup>以及客户平均中断持续时间指标(customer average interruption duration index, CAIDI)进行量化<sup>[14]</sup>。此外,故障期间配电网的平均削减功率(average curtailed power, ACP)反映了系统在故障或停电情况下的负荷损失程度<sup>[15]</sup>。因此,ACP也可作为量化配电网可靠性的指标。

配电网重构过程中考虑的约束条件主要有配电网功率平衡约束、配电网电压幅值约束、配电网支路功率约束、配电网线路最大载流量约束、开关次数约束以及DG出力约束等。此外,配电网辐射状约束限制了网络的最大负载能力与节点连接性。因此,在求解ADNDR策略时还需考虑辐射状约束。配电网辐射状约束的建模方法主要有以下5种:基于潮流约束的建模方法、基于虚拟潮流的建模方法、基于图的生成树的建模方法、基于供电路径的建模方法以及基于供电环路的描述方法。其中,最为常用的为基于供电环路的建模方法,该方法需要满足2条约束:(1)网络中有 $N-N_s$ 条闭合的支路,其表达式如式(1)所示;(2)配电网中不存在连通的供电环路,其表达式如式(2)所示<sup>[16]</sup>。

$$\sum_{b=1}^B x_{b,t} = N - N_s \quad (1)$$

$$\sum_{m=1}^{M_l} x_{lm,t} \leq M_l - 1 \quad (2)$$

式中: $B$ 为配电网中的支路总数; $b$ 为支路序号,且 $b=1, 2, \dots, B$ ;  $x_{b,t}$ 为 $t$ 时刻第 $b$ 条支路的状态,当支路闭合时 $x_{b,t}=1$ ,当支路断开时 $x_{b,t}=0$ ;  $N$ 为网络中的节点数量; $N_s$ 为网络中的电源数量; $M_l$ 为第 $l$ 个供电环路中的支路数量, $l=1, 2, \dots, L$ ;  $L$ 为网络中供电环路总数; $x_{lm,t}$ 为 $t$ 时刻第 $l$ 个供电环路中的第 $m$ 条支路的状态,当支路闭合时 $x_{lm,t}=1$ ,当支路断开时 $x_{lm,t}=0$ 。

综上所述,目前ADNDR数学模型中的优化目标与约束条件已逐步多元化,这会造成ADNDR的数学模型变得更加复杂,导致求解过程需要更长的计算时间,进而限制模型的实时性,使得动态重构的决策过程变得缓慢或难以实施。此外,随着模型复杂性的提升,求解算法需要处理的数据量也相应增大,这会增大求解过程中的计算误差,进而影响模型求解的准确性和可靠性。

## 2) 数学模型精细化

为了更准确地描述主动配电网结构的复杂性,ADNDR数学模型必须足够精细,以捕捉各种因素

之间的相互作用和影响。文献[17]考虑了电阻随温度的变化情况;文献[18]认为导体寿命受温度影响,同时由于电流大小决定导体温度,在制定故障重构策略时考虑了故障率与电流大小的关系;文献[19]考虑了碳排放交易机制对主动配电网经济调度的影响;文献[20]建立了极端冰雪天气情况下的ADNDR数学模型;文献[21]考虑了储能系统(energy storage system, ESS)状态的耦合特性;文献[22]考虑了ESS的循环老化成本。

综上所述,目前已有越来越多的要素被纳入ADNDR数学模型的考虑范围之中,这也会使得模型的复杂度不断上升,导致重构策略的求解愈发困难。此外,过度精细化的模型可能过于依赖特定的数据或场景,导致其在面对不同情况时泛化能力受限,进而限制了模型在实际应用中的灵活性和适应性。

## 3) 不确定性因素处理手段多样化

随着DG的广泛接入和负荷需求的日益多样化,负荷与DG出力的不确定性给ADNDR的数学模型构建带来了前所未有的挑战。这种不确定性不仅增加了模型构建的复杂性,还可能影响优化结果的准确性和可靠性。目前,ADNDR策略研究在构建数学模型时,以各种不同的处理手段将这些不确定性纳入考虑范围。本文从不确定性因素、处理方式及不足之处3个方面对其进行分析,如表1所示<sup>[23-30]</sup>。

文献[23]将DG等效为“负”的负荷并统一转化为PQ节点,这种做法忽视了不同DG模型之间存在的差异性。文献[24]根据DG并网方式的差异将其分为不同类型的节点,再转化为PQ节点进行数学模型构建。上述文献均忽视了DG出力的不确定性。为了有效考虑DG出力的不确定性,部分文献采用概率分布模型,该模型不仅提供了DG出力的期望值,还给出了其可能的波动范围和概率分布,从而提高对DG出力预测的准确性。文献[25]采用威布尔分布描述DG出力的不确定性,采用正态分布描述负荷的不确定性。文献[26]采用威布尔分布描述风力发电机出力的不确定性,采用贝塔分布描述光伏发电出力的不确定性,采用高斯分布描述电动汽车(electric vehicle, EV)充电负荷的不确定性。

然而,由于概率分布函数不易确定,而且求解过程中的卷积计算比较复杂,因此这种方式存在一定局限性。场景分析法通过生成具体的场景和相应的概率来描述不确定性,避免了建立复杂的数学模型。文献[27]以Wasserstein概率距离指标对DG出

表 1 不确定性因素处理方式分析

Table 1 Analysis of handling approaches for uncertainty factors

不确定性因素	处理方式	不足之处
DG <sup>[23]</sup>	等效为 PQ 节点	未考虑 DG 出力的不确定性
DG <sup>[24]</sup>	将不同 DG 视作不同节点	建模与计算过程较为复杂
DG、负荷 <sup>[25-26]</sup>	使用概率分布模型描述	需要大量的历史数据
DG <sup>[27]</sup>	对 DG 出力进行场景划分	生成可能的场景
DG <sup>[28]</sup>	对 DG 出力进行场景划分	准确性不足
DG、负荷 <sup>[29-30]</sup>	使用区间数描述波动范围	

力场景进行划分,提出基于场景划分的配电网重构方法。文献[28]将风速与风机出力之间的关系划分为 3 个典型场景:零、欠额定和额定输出场景。文献[29-30]利用区间数描述 EV 充电负荷的不确定性、DG 出力和负荷容量预测的不确定性,提出了各种不确定性量的区间描述方法。该方法虽然同样避免了建立复杂的数学模型,但其以范围形式估计负荷与 DG 出力,在精确度方面存在一定局限性。

综上所述,传统方法在处理不确定性因素方面均呈现出一定的局限性,这凸显了对于高效处理不确定性因素算法的迫切需求。为了推动 ADNDR 策略研究迈向更深层次,亟需一种能够有效适应这些不确定性因素的算法对 ADNDR 策略进行求解。

## 2 主动配电网动态重构编码方式

主动配电网包含大量的节点和线路,若直接对配电网进行重构,则算法的解空间维度较高。以 IEEE 33 节点电力系统为例,其示意图如附录 A 图 A1 所示。若不对配电网进行编码,则系统中的 14 个可动作开关均存在断开与闭合 2 种选择,其解空间的维度将达到  $2^{14}=16\ 384$  维。将领域知识和先验信息进行整合,对配电网进行编码,以结构化开关状态的解空间,从而使搜索算法能更高效地在其中进行搜索,以提高算法的求解效率。目前,最常见的 ADNDR 编码方式有 3 种:基于开关(支路)的编码方式、基于支路组的编码方式和基于基本回路的编码方式。

### 2.1 基于开关(支路)的编码方式

目前常见的基于开关的编码方式有二进制开关编码<sup>[31]</sup>与十进制开关编码<sup>[32]</sup>。二进制开关编码对配电网中位于环路内的每个开关的状态使用 0 或 1 进行编码,其中,0 表示开关断开,1 表示开关闭合。由于不在环路中的开关以及直接与电源相连的开关

必须闭合,故不对其进行编码。以 IEEE 33 节点电力系统为例,除与电源直接相连的 1 号开关外,该系统共 13 个开关可进行动作,因此对该系统的开关进行二进制编码所得到的解空间维度为  $2^{13}=8\ 192$  维,但其中存在大量的不可行解,且每个动作需要 1 个 13 维向量进行存储。

十进制开关编码对每个基本回路中的开关从 1 到  $N_{sw}$  进行编码,其中  $N_{sw}$  为基本回路中的可动作开关数量。之后,按顺序在每个基本回路中选择 1 个开关断开,并且相邻基本回路的公共支路中只能断开 1 个开关。以 IEEE 33 节点电力系统为例,其十进制开关编码表如附录 A 表 A1 所示。首先,该电力系统有 5 个基本回路,因此其动作  $X$  为断开 5 个开关  $\{x_1, x_2, x_3, x_4, x_5\}$ , 且有:

$$\begin{cases} 1 \leq x_1, x_4, x_5 \leq 4 \\ 1 \leq x_2 \leq 6 \\ 1 \leq x_3 \leq 3 \end{cases} \quad (3)$$

其次,该电力系统中的 1 号基本回路与 2 号基本回路中均含有支路 27,因此当 1 号基本回路断开支路 27 时,2 号基本回路只能在除支路 27 外的剩余 5 条支路中选择 1 条支路断开。若只考虑 1 个基本回路断开 1 个开关,则有  $4 \times 6 \times 3 \times 4 \times 4 = 1\ 152$  种可行解,该编码方式可在此基础上继续降低动作空间,大大减少了重构过程中的计算复杂度。文献[33]使用该编码方式降低重构过程中的解空间维度,但可行解的数量也因此减少。文献[34]结合该编码方式与遗传算法,大大降低了故障重构策略的求解时间。

### 2.2 基于支路组的编码方式

基于支路组的编码方式将重构效果相同的支路合并成 1 个支路组,从而获得网络同胚图,并对其进行编码。再根据“若配电网存在  $N_F$  个基本回路,当且仅当断开属于  $N_F$  个不同支路组的  $N_F$  条支路时,该网络才可能满足辐射状约束”这一约束条件来确定解空间,可以进一步降低解空间的维度<sup>[35]</sup>。以 IEEE 33 节点系统为例,首先删去系统中不构成回路的支路,得到系统的简化图,如附录 A 图 A2(a)所示;然后删去简化图中度为 2 的节点,得到系统的同胚图,如附录 A 图 A2(b)所示;最后对同胚图中的支路进行编码,得到附录 A 图 A2(c)中基于支路组的编码图。由附录 A 图 A2(c)可知,该网络有 12 条支路组,又由于该网络有 5 个基本回路,因此使用基于支路组的编码方式对其进行编码,可将解



文献[43-44]将 ADNDR 数学模型转化为 MICP 问题进行求解。文献[43]首次将 ADNDR 与无功补偿协同优化数学模型转化为 MICP 问题；文献[44]通过在数学模型中添加新的辐射状约束条件的方式，使其适用于不同规模的配电网。文献[45]使用改进 distFlow 模型将 ADNDR 与无功协调优化数学模型转化为 MIQP 问题，取得了更好的优化效果。文献 [46-47] 使用大 M 法与二阶锥松弛法 (second-order cone relaxation, SOCR) 将 ADNDR 数学模型转化为 MISOCP 问题进行求解。

上述文献通过松弛法将 ADNDR 数学模型转化为 MILP、MIQP、MIQCP、MICP 及 MISOCP 进行求解。数学优化算法可以有效解决搜索空间较小的优化问题，保证其收敛至最优解。因此，数学优化算法在处理小规模配电网重构问题时能够找到最优的重构策略，但数学优化算法在处理高维混合整数非线性优化问题时，存在着数学模型构建复杂度高、计算时间长以及松弛误差大等不足之处。

### 3.2 启发式算法

启发式算法是根据电力系统操作员的经验与问题的特殊特征而提出的一种最优重构策略算法[48]。相比于传统的数学优化算法，启发式算法通常具有更高的求解效率，能够在较短的时间内找到较优或接近最优的解决方案。下文从优化算法、优化目标、创新点及不足之处 4 个方面对基于启发式算法的 ADNDR 相关文献进行分析，如表 3 所示[49-57]。

目前常见的启发式算法分别是支路交换 (branch exchange, BE) 算法与最优流模式 (optimal power flow pattern, OPFP) 算法。文献[49-51]使用支路交换 BE 算法对配电网进行重构。文献[49]所提策略在优化过程中可以直接满足配电网辐射状约束，无需其他附加准则，但其对于配电网初始的拓扑结构的依赖性极高。文献[52-53]使用 OPFP 算法对配电网进行重构。文献[52]在初始时刻将开关全部闭合，在重构过程中无需初始配电网的开关状态，但由于回路间的相互影响，该策略仍容易陷入局部最优解。在文献[52]的基础上，文献[53]推导出了一种新的全局 OPFP。

综上所述，基于 BE 算法与 OPFP 算法的 ADNDR 策略均存在各自的弊端，文献[54-57]使用改进的启发式算法以弥补上述启发式算法在解决 ADNDR 问题时的不足。文献[54]提出了一种开关打开和交换 (switch opening and exchange, SOE) 的启发

表 2 基于数学优化算法的配电网重构策略分析

Table 2 Analysis of distribution network reconfiguration strategies based on mathematical optimization algorithms

优化算法	松弛方法	创新点	不足之处
MILP <sup>[39]</sup>	大 M 法	同时制定最优 ADNDR 策略和最佳线路类型选择策略	M 值选择不当会严重影响所求策略的有效性
MILP <sup>[40]</sup>	PL 法	得到了一种考虑电压稳定性的主动配电网重构最优策略	PL 法会导致求解问题的计算复杂度增加
MIQCP <sup>[41]</sup>	DCHR 法	提出了一种更为紧密的松弛方法	建模需要精确的配电网参数
MIQCP <sup>[42]</sup>	大 M 法	提出了 2 种混合整数凸优化模型并对其进行了比较	根据二次约束编写非线性潮流方程复杂度高
MICP <sup>[43]</sup>	大 M 法、多面体近似	首次将 ADNDR 与无功优化协调模型构建为 MICP	引入辅助变量会增加问题的复杂度
MICP <sup>[44]</sup>	锥松弛法	所提模型可在极短时间内精确求解最优重构策略	未考虑节点 $i$ 与节点 $j$ 之间的相角
MIQP <sup>[45]</sup>	小角正弦近似法	将改进的 distFlow 模型应用于 ADNDR 与无功协调优化问题	小角正弦近似仍存在一定误差
MISOCP <sup>[46]</sup>	大 M 法、SOCR	所提模型适用于不同规模配电网	建模需要精确的配电网参数
MISOCP <sup>[47]</sup>	大 M 法、SOCR	提出了一种协调配电网重构与孤岛分区的故障恢复策略	建模需要精确的配电网参数

表 3 基于启发式算法的主动配电网动态重构策略分析

Table 3 Analysis of active distribution network dynamic reconfiguration strategies based on heuristic algorithms

优化算法	优化目标	创新点	不足之处
BE <sup>[49]</sup>	网络损耗	考虑了 DG 接入配电网后的双向潮流对 BE 的影响	对初始配电网拓扑结构依赖性极高
BE <sup>[50]</sup>	电能质量	提出了 3 种启发式算法对 ADNDR 问题进行求解	对初始配电网拓扑结构依赖性极高
BE <sup>[51]</sup>	网络损耗、电能质量	同时改善配电网的网络损耗和电能质量	对初始配电网拓扑结构依赖性极高
OPFP <sup>[52]</sup>	网络损耗	重构时不依赖于初始配电网拓扑结构	无法避免回路间的相互影响
OPFP <sup>[53]</sup>	网络损耗	新的全局 OPFP 为重构策略提供最佳起点和收敛方向	对初始配电网拓扑结构依赖性极高
SOE <sup>[54]</sup>	网络损耗	所提算法结合了 BE 算法与 OPFP 算法的优点	算法泛化性能较差
SOE <sup>[55]</sup>	网络损耗	将启发式算法与 MISOCP 算法进行了对比	算法需要准确的配电网参数进行模型构建
FO <sup>[56]</sup>	网络损耗	提出了一种新的启发式算法以解决 ADNDR 问题	解决低维 ADNDR 问题时效率低于数学优化算法
MST <sup>[57]</sup>	网络损耗	将 MST 算法与启发式规则进行结合	容易陷入局部最优解

式算法以解决 ADNDR 问题, 该算法结合了 BE 算法与 OPFP 算法的优势。文献[55]将该算法与 MISOCP 进行比较, 结果表明该算法具有更好的优化效果与求解效率。文献[56]提出了一种修复和优化(fix and optimize, FO)的启发式算法以解决 ADNDR 问题。文献[57]将最小生成树(minimum spanning tree, MST)算法与改进启发式规则相结合, 可在不依赖于配电网初始拓扑结构的前提下进行重构。

上述文献采用 BE、OPFO、SOE、FO 以及 MST 算法等启发式算法对 ADNDR 问题进行求解。由优化结果可得, 上述算法有着优于传统数学算法的求解效率, 但仍存在需要精确的配电网参数、极易陷入局部最优解以及泛化性能较差等不足。

### 3.3 元启发式算法

元启发式算法是一种使用学习策略、探索和利用搜索空间寻找最优或接近最优解的迭代过程算法。下文从优化算法、优化目标、创新点及不足之处 4 个方面对基于元启发式算法的 ADNDR 策略相关文献进行了分析, 如表 4 所示<sup>[58-67]</sup>。

文献[58-59]基于改进遗传算法(genetic algorithm, GA)对配电网进行重构。改进 GA 使用基于种群的过程生成初始种群, 提高了 ADNDR 的求解效率。文献[60-61]基于粒子群优化(particle swarm optimization, PSO)算法对主动配电网进行重构。文献[61]提出了一种时变加速系数辅助的二进制 PSO 算法对 ADNDR 问题进行求解, 降低了粒子群中粒子陷入局部最优的概率。但 PSO 算法主要依赖于局部信息和邻域搜索, 因此该算法极易陷入局部最优解。文献[62-63]基于模拟退火(simulated annealing, SA)算法对主动配电网进行重构。文献[62]提出了基于 SA 算法的故障恢复重构策略。在此基础上, 文献[63]将 MST 算法用于求解 ADNDR, SA 算法用于求解无功补偿策略。文献[64-65]基于帝国主义竞争算法(imperialist competitive algorithm, ICA)对主动配电网进行重构。文献[66-67]基于和声搜索(harmony search, HS)算法对配电网进行重构。文献[67]在文献[66]的基础上增加了路径重新链接(path relinking, PR)阶段, 加速了算法的收敛。该策略在大规模配电网中的重构效果优于文献[66]所提策略。

上述文献采用 GA、PSO、SA、ICA 以及 HS 算法等元启发式算法对 ADNDR 问题进行求解。其相较于数学优化算法具有更高的求解效率, 相较于启发式算法具有更优的全局搜索能力及泛化性能,

表 4 基于元启发式算法的主动配电网动态重构策略分析  
Table 4 Analysis of active distribution network dynamic re-configuration strategies based on metaheuristic algorithms

优化算法	优化目标	创新点	不足之处
GA <sup>[58]</sup>	网络损耗、LB	使用 BE 生成初始种群, 避免了不可行解的产生	未考虑可再生能源接入
GA <sup>[59]</sup>	开关成本、网络损耗、可靠性	使用基于种群的过程生成初始种群	求解高维 ADNDR 问题效率较低
PSO <sup>[60]</sup>	网络损耗、电压偏移	使用混合 PSO 算法求解 ADNDR 与 DG 的选址定容	容易陷入局部最优解
PSO <sup>[61]</sup>	网络损耗	提出了一种新的时变加速系数辅助二进制 PSO 算法	建模需要精确的配电网参数
SA <sup>[62]</sup>	网络损耗、故障恢复时间等	使用 SA 算法求解 ADNDR 与无功协调优化策略	求解时间极长
SA <sup>[63]</sup>	用电成本、故障恢复时间等	在文献[62]的基础上增加了 MST 算法作为辅助	无法应对高维 ADNDR 问题
ICA <sup>[64]</sup>	网络损耗	同时进行 ADNDR、DG 选型与调压变压器档位选择	参数调节较困难
ICA <sup>[65]</sup>	网络损耗、电压偏移	使用牛顿-拉夫逊法计算潮流	容易陷入局部最优解
HS <sup>[66]</sup>	网络损耗	在音高调整过程中获得由数学方程建立的自适应行为	容易陷入局部最优解
HS <sup>[67]</sup>	网络损耗	在文献[66]所提策略的基础上增加了 PR 阶段	参数调节较困难

但仍存在需要精确的配电网参数、算法参数调节较为困难以及较容易陷入局部最优等不足。

### 3.4 传统机器学习算法

机器学习算法是一种用于训练计算机系统从数据中进行学习的通用方法。该类方法通过利用大量历史数据进行学习, 可以更好地理解配电网的运行特征和规律。由于该类方法能够根据问题的复杂性和变化自动调整参数和搜索策略, 因此具有较强的自适应性, 从而能够更好地适应不同规模的 ADNDR 问题。下文将从优化算法、优化目标、创新点以及不足之处 4 个方面对基于传统机器学习算法的 ADNDR 策略相关文献进行分析, 如表 5 所示<sup>[68-79]</sup>。

文献[68-69]使用分类回归树(classification and regression tree, CART)算法与 K 近邻(K-nearest neighbors, KNN)算法求解 ADNDR 策略。文献[70-71]使用 K 均值(K-means, KM)聚类算法与模糊 C 均值(fuzzy c-means, FCM)聚类算法求解 ADNDR 策略。人工神经网络(artificial neural network, ANN)通过模拟生物神经网络的工作原理, 能够较好地拟合 ADNDR 中各决策变量间的非线性关系。文献[72-73]使用改进的 ANN 算法求解 ADNDR 策略。

表 5 基于传统机器学习算法的主动配电网动态重构策略分析

Table 5 Analysis of active distribution network dynamic reconfiguration strategies based on traditional machine learning algorithms

方向	优化算法	优化目标	创新点	不足之处
传统监督学习算法	CART <sup>[68]</sup>	网络损耗	考虑了配电网运行数据与重构方案之间的关系	处理高维 ADNDR 问题效果不佳
	CART <sup>[69]</sup>	电压稳定性	将 KNN 与 CART 结合以提高重构后的电压稳定性	处理高维 ADNDR 问题效果不佳
无监督学习算法	KM 聚类 <sup>[70]</sup>	网络损耗	采用 KM 聚类算法提取每个季度代表性负荷情景	KM 聚类对模糊性强的数据处理能力弱
	FCM 聚类 <sup>[71]</sup>	网络损耗	提出了一种基于 FCM 聚类算法的 ADNDR 策略	处理高维 ADNDR 问题效果不佳
人工神经网络和深度神经网络算法	ANN <sup>[72]</sup>	网络损耗	考虑了不同消费模式的负荷曲线对重构策略的影响	神经网络规模有待降低
	ANN <sup>[73]</sup>	网络损耗	每个 ANN 对应一个开关, 进一步降低 ANN 规模	ANN 对环境的感知和学习能力不足
	CNN <sup>[74]</sup>	网络损耗、LB、电压偏移	在不依赖时域模拟的前提下进行优化	未考虑可再生能源出力的随机性
	CNN <sup>[75]</sup>	网络损耗、LB、电压偏移	构建 CNN 模型预测电压稳定性指数	CNN 处理图结构与时序数据的能力较弱
	LSTM <sup>[76]</sup>	网络损耗、电压偏移、LB	利用 LSTM 准确预测各节点在不同时期的负荷	未考虑配电网三相不平衡性
	LSTM <sup>[77]</sup>	网络损耗、LB	可预测 DG 出力 and 负载的精确概率分布	建模需要精确的配电网参数
	GCN <sup>[78]</sup>	网络损耗、电压偏移	考虑了节点之间的连接关系	建模需要精确的配电网参数
GCN <sup>[79]</sup>	供电恢复速度	引入线图空间转换方法对配电网拓扑进行空间变换	建模需要精确的配电网参数	

深度神经网络(deep neural network, DNN)在 ANN 的基础上增加了隐藏层的数量, 以获得更好的环境感知能力与学习能力。常用于处理 ADNDR 的 DNN 算法有卷积神经网络(convolution neural network, CNN)算法、长短期记忆网络(long short-term memory, LSTM)网络算法与图卷积网络(graph convolutional network, GCN)算法<sup>[74-79]</sup>。文献<sup>[76-77]</sup>基于 LSTM 进行 ADNDR 策略研究。配电网数据是具有图结构的数据, 使用 DNN 对其进行处理存在局限性, 而 GCN 善于处理具有图结构的数据。文献<sup>[78-79]</sup>基于 GCN 进行 ADNDR 策略研究。文献<sup>[79]</sup>利用 GCN 极强的非线性拟合能力, 建立配电网快速供电恢复智能决策模型, 从而满足了高比例 DG 接入下配电网的供电恢复需求。

上述文献采用 CART、KNN、聚类、ANN 以及 DNN 等算法对 ADNDR 问题进行求解。上述算法均采用高效的搜索策略与并行化计算, 因此相较于数学优化算法具有更高的求解效率。此外, 传统机器学习算法采用全局搜索策略与随机性搜索策略搜索最优解, 因此相较于启发式算法与元启发式算法, 传统机器学习算法具有更优的全局搜索能力, 但仍存在处理高维 ADNDR 问题效果不佳以及需要确定配电网参数等不足之处。

综上所述, ADNDR 问题作为一个复杂的高维混合整数非线性随机优化问题, 基于传统算法的 ADNDR 策略存在着各自的优势与不足, 如表 6 所示<sup>[39-79]</sup>。由表 6 可知, 基于传统算法的 ADNDR 策略在部分方面具有一定优势, 但仍存在容易陷入局部最优、需要显式划分重构时段、对于不确定因素

的适应性差、需要完整的配电网参数进行模型构建以及不适用于大规模配电网等不足之处。因此, 亟需一种兼顾求解精度、求解效率、泛化性能以及全局搜索能力的算法对 ADNDR 进行快速且准确地求解。

#### 4 基于深度强化学习算法的重构策略

ADNDR 数学模型中包含了多重不确定性因素, 如 DG 出力的不确定性和负荷的不确定性。因此, 该数学模型本质是一个高维、非线性随机优化数学模型。传统数学算法在解决该问题的过程中存在计算速度慢、无法有效处理大规模随机变量等问题, 而启发式算法容易陷入局部最优解。为了更加快速、准确地对 ADNDR 数学模型进行求解, 现有研究引入了强化学习算法。强化学习算法的详细介绍见附录 B。

由附录 B 可知, DRL 算法在求解 ADNDR 策略时能够兼顾不确定因素适应性、求解精度、求解效率、泛化性能与全局搜索能力。因此, 许多研究采用 DRL 算法求解 ADNDR 策略, 从而对配电网的经济性与可靠性进行优化。

##### 4.1 基于 DRL 的配电网经济性优化

由第 1 章可知, ADNDR 策略主要通过改善网络损耗、负荷均衡度、开关操作次数、综合费用以及 SACP 的方式对配电网的经济性进行优化。文献<sup>[38,80-81]</sup>使用基于价值的 DRL 算法进行 ADNDR 策略研究。文献<sup>[80]</sup>根据网络损耗与违反约束条件的惩罚设置奖励函数, 并从历史运营数据集中生成可靠的综合运营数据, 与训练过程中生成的经验数

表6 传统主动配电网动态重构算法优劣势分析

Table 6 Analysis of the advantages and disadvantages of traditional active distribution network dynamic reconfiguration algorithms

算法类别	具体算法	求解精度	求解效率	泛化性能	全局搜索能力	算法优势	算法不足之处
数学优化算法	MILP <sup>[39-40]</sup>					能精确找到松弛后的最优解	对 ADNDR 进行松弛会产生一定误差
	MIQCP <sup>[41-42]</sup>					可以保留二次约束	根据二次约束编写潮流方程极为复杂
	MICP <sup>[43-44]</sup>	高	低	弱	优	能精确找到松弛后的最优解	无法处理相角
	MIQP <sup>[45]</sup>					可以处理相角	无法预测负荷与 DG 出力
	MISOCP <sup>[46-47]</sup>					能精确找到小规模配电网的最优重构策略	不适用于大规模配电网
启发式算法	BE <sup>[49-51]</sup>					可以在优化过程中直接满足辐射状约束	对初始配电网结构依赖程度高
	OPFP <sup>[52-53]</sup>					重构时不依赖于初始配电网拓扑结构	无法避免回路间的相互影响
	SOE <sup>[54-55]</sup>	低	较低	较弱	弱	既不依赖初始结构又能避免回路间的相互影响	泛化性能较差
	FO <sup>[56]</sup>					在大规模配电网中求解效率优于数学优化算法	在小规模配电网中求解效率低于数学优化算法
	MST <sup>[57]</sup>					重构时不依赖于初始配电网拓扑结构	容易陷入局部最优解
元启发式算法	GA <sup>[58-59]</sup>					可以适应非线性混合整数优化问题	计算效率随着染色体长度的增加而降低
	PSO <sup>[60-61]</sup>					需要调整的参数很少	极易收敛到局部最优解
	SA <sup>[62-63]</sup>	较低	较高	较优	较弱	算法易于实现	求解结果取决于初始的配电网结构
	ICA <sup>[64-65]</sup>					不易收敛于局部最优解	参数调节较困难
	HS <sup>[66-67]</sup>					需要调整的参数极少	容易陷入局部最优解
传统机器学习算法	CART <sup>[68-69]</sup>					算法可解释性强	在大规模配电网中效果不佳
	聚类 <sup>[70-71]</sup>					可以对高维数据进行降维	计算复杂度高
	ANN <sup>[72-73]</sup>	较高	高	优	较优	可以处理复杂的非线性数据	ANN 对环境的感知和学习能力不足
	CNN <sup>[74-75]</sup>					特征提取能力强	CNN 处理图结构与时序数据的能力较弱
	LSTM <sup>[76-77]</sup>					可以很好地处理时序数据	需要完整的配电网参数进行模型构建
GCN <sup>[78-79]</sup>					可以很好地处理图结构数据	参数调节较困难	

据一起存入经验回放池, 用于之后的训练过程, 进而有效改善了当前 DRL 算法在样本性能方面的局限性, 但该方法没有考虑开关的动作次数限制。在文献[80]的基础上, 文献[81]增加了对于开关动作次数的限制, 并提出了一种更新动作空间的方法。该方法在筛选过程中, 将那些动作后不满足开关动作次数限制的开关从动作空间中剔除, 以此确定动作空间的子集。随后, 智能体仅需在精简后的动作空间中选择重构开关, 大大降低了算法复杂度。文献[38]通过基于基本回路的编码方式以满足配电网的辐射状约束, 大大降低了动作空间的维度, 并通过添加网络权重的扰动, 可以在不调整探索参数的情况下实现自动探索, 从而加速训练过程, 以提高优化性能。

相较于基于价值的 DRL 算法, 价值和策略结合的 DRL 算法有着更好的效果。文献[82-83]基于价值和策略结合的 DRL 算法对主动配电网进行重构。文献[82]使用学习到的策略与历史运营策略之间的 KL 散度正则化对奖励函数进行规范, 以降低直接应用 DRL 算法所产生的外推误差。文献[83]通过预学习配电网运行历史中网络损耗最低的开关状态,

以快速获得不同电源和负载状态下配电网的最优开关动作。同时, 其通过蒙特卡洛树搜索(Monte Carlo tree search, MCTS)算法进行辅助决策, 当 DRL 算法预学习的重构策略不满足约束条件时, MCTS 可以通过备份方案实现快速重构。文献[84-85]基于 MADRL 算法进行 ADNDR 策略研究。文献[84]为配电网中的每个最小回路配置 1 个智能体, 每个智能体通过 DDPG 算法求解其所属回路的最优重构开关。在训练期间, 评论家网络从全局环境中学习; 而在执行过程中, 每个智能体仅观察所属回路的状态, 并从演员网络输出自己的策略。这种方法避免了将 DDPG 算法直接引入多智能体环境时, 可能出现的策略梯度方差大和不平稳性问题。文献[85]提出了一种双时间尺度 ADNDR 策略。该策略先将 ADNDR 数学模型松弛为一个 MISOCP 问题, 并在日前预求解最优重构策略。在日内, 该策略将主动配电网分为多个区域, 并为每一个区域配置一个智能体, 每个智能体使用双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法求解该区域内的新能源出力削减量; 该文献将参数共享机制应用于 MATD3 算法, 以克服多智能

体环境的不平稳性, 保证其在解决不同规模的 ADNDR 问题时的适用性。上述文献均仅将所提策略与基于传统算法或少数基于 DRL 算法的 ADNDR 策略进行比较, 因此无法得知各 DRL 算法的适用场景。文献[86]将 5 种不同 DRL 算法用于解决 ADNDR 问题, 并对优化结果进行了比较, 得到软演员评论家算法的收敛速度最快、而 DQN 算法的稳定性和有效性最优的结论。

传统联络开关使用机械部件进行连接和断开, 长期使用容易导致磨损, 需要定期维护和更换。此外, 其在进行连接和断开操作时, 需要较长时间, 因此无法满足一些对动作速度要求较高的应用场景。相比之下, 软开关(soft open point, SOP)通常具有更快速、更精确的动作特性, 且无需定期维护。因此, 在配电网中使用 SOP 代替传统联络开关具有多方面的优势。文献[87-88]考虑了基于 DRL 算法的含 SOP 配电网重构策略。文献[87]构建了基于优先经验回放(prioritized experience replay, PER)的 DQN 算法的 SOP 与 ADNDR 的协同优化策略。该策略通过调动源网侧灵活资源, 协同优化含 SOP 的可控有源设备与网络动态重构, 以降低配电网的网络损耗。在此基础上, 文献[88]使用 DDQN 算法求解重构策略, 使用 AC 算法对三相 SOP 的有功功率与无功功率进行控制, 从而更全面地分析电网的平衡性, 确保在动态重构过程中维持电网的平衡。

文献[89]提出了一种基于图胶囊神经网络(capsule graph neural network, CapsGNN)-PPO 算法的配电网重构策略。通过胶囊结构与图神经网络更好地表示节点之间的部分和整体关系, 从而更准确地学习节点之间的复杂连接模式。文献[90]综合考虑 SACP、电压偏移、网络损耗、电压稳定性和发电成本 5 个优化目标, 并为每个优化目标配置 1 个 DQN 智能体。多个智能体共同指导多目标细菌觅食优化(multi-objective bacterial foraging optimization, MBFO)算法求解重构策略, 以获得帕累托最优解, 由此平衡可再生能源的容纳量与配电网的运行状态。

#### 4.2 基于 DRL 的配电网可靠性优化

由第 1 章可知, ADNDR 策略主要通过减少故障发生的时间与频率或功率削减的方式对配电网可靠性进行优化。文献[91-93]使用基于价值的 DRL 算法进行 ADNDR 策略研究。文献[91]在满足全节点遍历约束后, 通过构建节点-支路入射矩阵的方式, 验证重构策略是否满足配电网的辐射状约束。

同时, 其以违反配电网约束的惩罚与 ACP 为奖励函数, 从而提高配电网运行可靠性, 但其未考虑故障发生的时间与频率。在文献[91]的基础上, 文献[92]提出了一种故障恢复重构策略, 该策略在提高配电网可靠性的同时, 优化了配电网的网络损耗。文献[93]结合配电网的网络损耗、电压偏移、故障率与越限惩罚设置奖励函数, 在优化配电网经济性的同时, 提升了配电网的可靠性; 此外, 该文献使用 2 点估计法(point estimate method, PEM)计算出更为准确的潮流分布, 确保了重构策略的准确性。上述策略虽能在一定程度上对配电网运行的可靠性进行优化, 但其均没有考虑配电系统中故障的随机性。文献[94]提出了一种 3 阶段主动配电网故障恢复重构策略, 该策略在第 1 阶段进行配电网日前优化调度, 在第 2 阶段使用 DDQN 算法求解最优故障重构策略, 在第 3 阶段使用 DDPG 算法确定所有 DG 的新设置点。该策略使得配电网可以确定任何故障所对应的最佳重构方案, 并更新 DG 的最佳设置点。

文献[95-96]将 LSTM 网络与指针网络嵌入 DRL 框架以考虑配电网重构过程中的时间序列数据。文献[95]结合越限惩罚与可靠性指标设置奖励函数, 并将 AC 算法中的全连接层改进为 LSTM。通过 LSTM 更好地捕捉状态序列之间的长期依赖关系, 以考虑状态序列之间转换的相互关联性, 提升配电网运行的可靠性。文献[96]结合故障恢复的快速性与有效性设置奖励函数, 利用改进指针网络结构配合 AC 算法实现适用于多类故障恢复策略的模型自学习和端到端计算; 此外, 其加入了动态信息作为辅助状态输入, 通过组合优化的方式对配电网可靠性进行了优化。文献[97]提出了一种基于 GCN-DQN 算法的主动配电网故障恢复重构策略以改善配电网的故障恢复率。该策略将 GCN 嵌入 DRL 框架, 并利用图数据表征配电网拓扑结构, 将其与配电网系统状态一同作为图神经网络的输入, 以此对图数据进行处理, 从而获得更优的动态重构策略。

上述文献采用 DQN、AC、DDPG 以及 PPO 等 DRL 算法对 ADNDR 问题进行求解。上述算法均在免模型的前提下, 利用 DRL 算法对于不确定因素的天然适应性进行重构策略求解, 因此, 基于 DRL 算法的 ADNDR 策略相较于传统算法具有更高的求解效率与更优的求解结果。表 7 从 DRL 算法、优化目标、创新点以及不足之处 4 个方面对基于 DRL 算法的主动配电网动态重构研究进行了总结梳理<sup>[38,80-97]</sup>。附录

表 7 基于深度强化学习算法的主动配电网动态重构策略

Table 7 Active distribution network dynamic reconfiguration strategy based on deep reinforcement learning algorithms

方向	DRL 算法	优化目标	创新点	不足之处
基于 DRL 的配电网经济性优化	DQN <sup>[80]</sup>	网络损耗	使用历史数据改善 DRL 在样本性能方面的局限性	没有考虑开关的动作次数限制
	DQN <sup>[81]</sup>	网络损耗、开关次数	更改动作空间以满足配电网重构的开关动作次数约束	基于支路的编码方式将导致动作空间极大
	DQN <sup>[38]</sup>	网络损耗、电压偏移	采用基于基本回路的编码方式、添加网络权重的扰动	DQN 收敛速度较慢
	AC <sup>[82]</sup>	网络损耗	通过 KL 散度降低使用 DRL 算法会引起的外推误差	算法参数根据经验选择
	AC <sup>[83]</sup>	网络损耗	预学习配电网运行历史中网络损耗最低的开关状态	历史数据无法包括所有运行状态
	MADDPG <sup>[84]</sup>	网络损耗、电压偏移	为配电网中的每个最小回路部署 1 个智能体	DDPG 在某些情况下容易陷入局部最优解
	MATD3 <sup>[85]</sup>	网络损耗	提出了一种双时间尺度 ADNDR 策略	TD3 对超参数敏感, 调参较困难
	DQN 等 <sup>[86]</sup>	网络损耗	将基于 5 种不同 DRL 算法的 ADNDR 策略进行比较	传统的联络开关需要定期维护和更换
	DQN <sup>[87]</sup>	网络损耗	调动源网侧灵活资源, 协同优化含 SOP 的主动配电网	未考虑电力系统三相不平衡性
	DDQN、AC <sup>[88]</sup>	运营成本	使用 AC 算法对三相 SOP 进行控制	传统的 DNN 无法充分考虑节点间的关联性
基于 DRL 的配电网可靠性优化	PPQ <sup>[89]</sup>	网络损耗	将 CapsGNN 嵌入 DRL 算法以读取配电网图数据	未考虑可再生能源消纳率
	DQN <sup>[90]</sup>	SACP、网络损耗等	采用 DRL 辅助 MBFO 算法以平衡多个优化目标	仍需具体的配电网参数进行模型构建
	DQN <sup>[91]</sup>	可靠性	使用节点-分支入射矩阵满足全节点遍历约束	未考虑网络损耗的优化
	DQN <sup>[92]</sup>	网络损耗、可靠性	使用 DRL 算法同时优化配电网的经济性与可靠性	未考虑可再生能源接入
	DQN <sup>[93]</sup>	网络损耗、可靠性	使用 2 m PEM 以计算出更准确的潮流分布情况	无法适应所有故障
	DDQN、DDPG <sup>[94]</sup>	运行成本、可靠性	可以确定系统中任何故障对应的最佳重构方法	没有考虑到状态序列切换之间的相互关联性
	AC <sup>[95]</sup>	网络损耗、可靠性	使用 LSTM 捕捉状态序列之间的长期依赖关系	无法充分考虑节点间的关联性
	AC <sup>[96]</sup>	可靠性	使用改进的指针网络辅助 DRL 算法	无法充分考虑节点间的关联性
	GCN-DQN <sup>[97]</sup>	故障恢复率	将 GCN 嵌入 DRL 算法以读取配电网图数据	没有考虑负荷的随机性

A 表 A2 对具体算法在求解主动配电网动态重构问题时的优点、缺点与改进方向进行了总结梳理。

## 5 总结与展望

### 5.1 未来研究方向

#### 5.1.1 多级动态重构研究

多级动态重构旨在通过动态调整配电网的多个级别或层次的结构和参数, 以优化配电网的性能和效率。通过对配电网的不同级别进行动态重构和优化, 可以实现更加精细化和智能化的控制, 从而更好地满足不断变化的负荷需求和运行条件。

多级动态重构涉及到对配电网的多个级别进行动态调整和改变, 这意味着它能够考虑整个系统的运行状态和性能, 从而实现更全面的系统性能提升。此外, 在主动配电网中, 由于 DG 出力的不确定性以及负荷侧需求侧响应的加入, 使得配电网的运行更加复杂。多级动态重构通过动态调整配电网的多个层级应对其不确定性, 从而提高配电网的稳定性。综上所述, 多级动态重构在主动配电网动态重构领域具有广泛的应用前景和重要的研究价值。

#### 5.1.2 主动配电网动态重构与无功补偿协同优化

无功补偿可以减少无功电流在电网中的流动, 从而降低网络的有功损耗, 且无功补偿能够改善电

网的电压水平, 减少电压波动和电压闪变等现象。此外, 在电网发生故障时, 无功补偿能够迅速响应, 提供必要的无功支撑, 从而维持系统的稳定性。然而, 目前大部分基于 DRL 算法的 ADNDR 策略研究均未使用无功补偿进行协同优化, 这将使得重构策略的优化效果受到限制, 无法充分发挥其潜力。因此, 在未来的研究中, 应考虑主动配电网动态重构与无功补偿协同优化。

#### 5.1.3 基于多目标强化学习算法的重构策略

从表 7 可以得知, 大多数基于 DRL 算法的 ADNDR 策略只对配电网经济性或可靠性进行优化。经济性优化通常会追求最低的运行成本, 这可能会导致配电网运营商实施减少备用设备和增加负荷率等措施, 进而降低了系统的可靠性。相反, 可靠性优化通常需要增加备用设备、提高系统冗余度等, 这会增加系统的运行成本。因此, 如何使得使用 DRL 算法求解的重构策略能够实现配电网的经济性与可靠性协同优化亟待考虑。

传统单目标 DRL 算法在优化多个目标时, 只能通过加权平均的方式将其转化为单目标优化问题, 这种方式在优化目标并非相互独立或相互矛盾时优化性能降低。多目标强化学习算法将 ADNDR 数学模型构建成为多目标马尔可夫决策过程

(multi-objective Markov decision process, MOMDP)。与传统的 MDP 不同, MOMDP 中的奖励函数  $R$  为矢量化的函数, 每一个优化目标均由  $R$  中的 1 个奖励信号  $r_n$  表示;  $\Omega$  为偏好空间, 由一系列偏好向量  $\omega$  组成。多目标强化学习算法在偏好空间  $\Omega$  中随机选择 1 组偏好向量  $\omega$  与状态  $s_t$  作为神经网络的输入, 以此更新动作价值函数。训练后的模型可以在多个优化目标之间进行权衡和协调, 使得算法找到 1 组 pareto 解, 以平衡所有的优化目标。

因此, 在未来的多目标 ADNDR 问题中, 可以将经济性优化目标与可靠性优化目标均由 1 个单一的奖励信号  $r$  表示, 并共同构成 MOMDP 中的奖励函数  $R$ 。在动作选择时, 智能体根据偏好向量  $\omega$  与奖励函数  $R$  的线性组合来选择最优重构策略, 以此对主动配电网经济性指标与可靠性指标进行协同优化。

#### 5.1.4 基于分层强化学习算法的重构策略

大规模主动配电网的动态重构问题具有极高的计算复杂度, 传统 DRL 算法在解决该问题时可能不够精确与高效。分层强化学习(hierarchical deep reinforcement learning, HDRL)算法可以将优化问题分解为多个子任务, 并在不同的层次上分别进行决策, 从而降低了优化问题的复杂性。此外, HDRL 算法可以通过将高层次的决策结果传递给低层次的决策模型, 实现信息的传递和共享, 从而提高学习效率和决策准确性。例如在 ADNDR 与无功协调优化问题中, 可以使用上层 DRL 算法确定离散动作(开关动作、可投切电容器档位), 使用下层 DRL 算法根据上层 DRL 算法的优化方案来确定连续动作(静止无功补偿器出力、DG 逆变器出力)以获得更好的优化策略。因此, 在未来研究中使用 HDRL 算法对 ADNDR 策略进行求解具有极高的研究价值。

## 5.2 结论

1) 在大规模配电网中, 使用传统方式满足辐射状约束产生的解空间维度极大, 基于开关的编码方式在一定程度上降低了解空间的维度, 但仍会产生大量的不可行解, 因此其性能并不理想。而基于支路组的编码方式虽进一步降低了不可行解出现的概率, 但其性能仍有待提升。基于基本回路的编码方式不仅可以有效降低动作空间的维度, 大大提升求解效率, 还可以保证其产生的动作均为可行解, 因此, 该编码方式是一种高效、可行的编码方式。

2) ADNDR 是一个高维非线性混合整数随机优化问题, 数学优化算法求解精度高, 全局搜索能力

较强但其求解效率低, 泛化性能弱, 启发式算法求解效率较高但容易陷入局部最优解且泛化性能较差, 元启发式算法泛化性能较好但仍容易陷入局部最优解; 传统机器学习算法求解效率高, 全局搜索能力强但其参数调节较为困难, 而通过 DRL 算法可以在免模型的前提下对重构策略进行实时优化, 克服了传统算法获取模型参数困难、算法实时性较差、需要显示划分重构时段以及无法有效处理高维的随机变量和非线性约束的弊端。

3) 基于 DRL 算法的 ADNDR 策略研究未来可考虑多级动态重构以及无功补偿协同优化等新型重构模型。此外, 还可以使用多目标深度强化学习算法与分层深度强化学习算法对 ADNDR 问题进行求解, 以获得更优的重构策略。

附录见本刊网络版(<http://hve.epri.sgcc.com.cn>)。

## 参考文献 References

- [1] 傅质馨, 李紫嫣, 朱俊澎, 等. “双碳”目标下需求侧管理机制研究综述及展望[J]. 电力信息与通信技术, 2023, 21(2): 1-12. FU Zhixin, LI Ziyan, ZHU Junpeng, et al. Overview and prospect of demand side management mechanism under “dual carbon” goal[J]. Electric Power Information and Communication Technology, 2023, 21(2): 1-12.
- [2] 吉兴全, 朱应业, 张玉敏, 等. 基于 DDQN-KRR 动态重构策略的综合能源系统运行优化[J]. 高电压技术, 2023, 49(8): 3195-3208. JI Xinquan, ZHU Yingye, ZHANG Yumin, et al. Optimization of comprehensive energy system operation based on DDQN-KRR dynamic reconstruction strategy[J]. High Voltage Engineering, 2023, 49(8): 3195-3208.
- [3] 张玉敏, 杨子震, 叶平峰, 等. 市场环境下考虑配网重构和需求响应的输配协同优化调度[J/OL]. 高电压技术, 2024: 1-15 [2024-12-28]. <https://doi.org/10.13336/j.1003-6520.hve.20240170>. ZHANG Yumin, YANG Zizhen, YE Pingfeng, et al. Coordinated optimization scheduling of transmission and distribution considering network reconstruction and demand response in a market environment[J/OL]. High Voltage Engineering, 2024: 1-15[2024-12-28]. <https://doi.org/10.13336/j.1003-6520.hve.20240170>.
- [4] MI Y, CHEN Y Y, YUAN M H, et al. Multi-timescale optimal dispatching strategy for coordinated source-grid-load-storage interaction in active distribution networks based on second-order cone planning[J]. Energies, 2023, 16(3): 1356.
- [5] 瞿合祚, 李晓明, 杨玲君, 等. 考虑负荷和分布式电源时变性的配电网多目标动态重构和 DG 调度[J]. 高电压技术, 2019, 45(3): 873-881. ZHAI Hezuo, LI Xiaoming, YANG Lingjun, et al. Multi objective dynamic reconstruction and DG scheduling of distribution networks considering the time-varying nature of loads and distributed power sources[J]. High Voltage Engineering, 2019, 45(3): 873-881.
- [6] SWAMINATHAN D, RAJAGOPALAN A, MONTROYA O D, et al. Distribution network reconfiguration based on hybrid golden flower algorithm for smart cities evolution[J]. Energies, 2023, 16(5): 2454.
- [7] 江昌旭, 刘晨曦, 林 铮, 等. 基于深度强化学习的电力系统暂态

- 稳定控制策略研究综述[J]. 高电压技术, 2023, 49(12): 5171-5186.
- JIANG Changxu, LIU Chenxi, LIN Zheng, et al. Review of power system transient stability control strategies based on deep reinforcement learning[J]. High Voltage Engineering, 2023, 49(12): 5171-5186.
- [8] 廖峰, 陈锦荣, 区伟潮, 等. 基于改进社会蜘蛛算法的有源配电网重构[J]. 电力系统及其自动化学报, 2023, 35(10): 125-133.
- LIAO Feng, CHEN Jinrong, OU Weichao, et al. Active distribution network reconfiguration based on improved social spider algorithm[J]. Proceedings of the CSU-EPSS, 2023, 35(10): 125-133.
- [9] 陈倩, 王维庆, 王海云. 基于含分布式电源的配电网优化运行研究[J]. 电测与仪表, 2023, 60(11): 19-28.
- CHEN Qian, WANG Weiqing, WANG Haiyun. Research on optimal operation of distribution network with distributed generation integration[J]. Electrical Measurement & Instrumentation, 2023, 60(11): 19-28.
- [10] 武晓朦, 李晨晨, 党博. 基于改进二进制 PSO 配电网动态重构[J]. 西安石油大学学报(自然科学版), 2024, 39(4): 124-131.
- WU Xiaomeng, LI Chenchen, DANG Bo. Dynamic reconfiguration of distribution network based on improved binary particle swarm optimization algorithm[J]. Journal of Xi'an Shiyou University (Natural Science Edition), 2024, 39(4): 124-131.
- [11] CHOBDARI M, SAMIEI MOGHADDAM M, DAVARZANI R, et al. Robust distribution networks reconfiguration considering the improvement of network resilience considering renewable energy resources[J]. Scientific Reports, 2024, 14(1): 23041.
- [12] SHAHEEN A, EL-SEHIEMY R, KAMEL S, et al. Optimal operational reliability and reconfiguration of electrical distribution network based on jellyfish search algorithm[J]. Energies, 2022, 15(19): 6994.
- [13] ANTENEH D, KHAN B. Reliability enhancement of distribution substation by using network reconfiguration a case study at debre berhan distribution substation[J]. International Journal of Economy, Energy and Environment, 2019, 4(2): 33-40.
- [14] MORADI-SARVESTANI S, DEHBOZORGI M R, RASTEGAR M. A three-stage reliability-centered framework for critical feeder identification, failure modes prioritization, and optimal maintenance strategy assignment in power distribution system[J]. Electric Power Systems Research, 2024, 230: 110215.
- [15] PENG Z Y, ZHANG W J, XU W C, et al. Conditional value at risk-based island partitioning and fault restoration reconfiguration of active distribution networks[J]. Frontiers in Energy Research, 2024, 12: 1460894.
- [16] 徐成司, 董树锋, 朱嘉麒, 等. 基于供电环路非连通条件的配电网辐射状约束描述方法[J]. 电力系统自动化, 2019, 43(20): 82-89.
- XU Chengsi, DONG Shufeng, ZHU Jiaqi, et al. Description method of radial constraints for distribution network based on disconnection condition of power supply loop[J]. Automation of Electric Power Systems, 2019, 43(20): 82-89.
- [17] 石泽轩, 张旭东, 应展烽, 等. 计及电缆温度约束的配电网重构研究[J]. 南京理工大学学报, 2019, 43(5): 563-570.
- SHI Zexuan, ZHANG Xudong, YING Zhanfeng, et al. Research on distribution network reconfiguration based on cable temperature constraint[J]. Journal of Nanjing University of Science and Technology, 2019, 43(5): 563-570.
- [18] KHASANOV M, KAMEL S, ABDUBANNAEV J, et al. Distribution network planning with DG units considering the network reconfiguration and reliability[J]. E3S Web of Conferences, 2023, 461: 01053.
- [19] 刘剑, 徐兵, 陈世龙, 等. 考虑碳排放交易和可再生能源时序出力重构的主动配电网最优经济调度[J/OL]. 现代电力, 2024: 1-12 [2024-11-24]. <https://doi.org/10.19725/j.cnki.1007-2322.2023.0377>.
- LIU Jian, XU Bing, CHEN Shilong, et al. Optimal economic dispatching of active distribution network based on carbon emission trading and sequential output reconstruction of renewable energy[J/OL]. Modern Electric Power, 2024: 1-12 [2024-11-24]. <https://doi.org/10.19725/j.cnki.1007-2322.2023.0377>.
- [20] 吉兴全, 臧祥宇, 张玉敏, 等. 极端冰雪天气下计及孤岛划分与融合的配电网故障恢复[J/OL]. 电力系统自动化, 2024: 1-15 [2024-11-24]. <http://doi.org/10.7500/AEPS20240302003>.
- JI Xingquan, ZANG Xiangyu, ZHANG Yumin, et al. Service restoration of distribution network under extreme weather of ice and snow considering islanding partition and combination[J/OL]. Automation of Electric Power Systems, 2024: 1-15 [2024-11-24]. <http://doi.org/10.7500/AEPS20240302003>.
- [21] 黄莘杰, 方斯顿, 罗颖冰, 等. 双向潮流下考虑储能系统状态耦合特性的配电网电压协调管理策略[J/OL]. 中国电机工程学报, 2024: 1-12 [2024-11-24]. <http://kns.cnki.net/kcms/detail/11.2107.tm.20240613.1305.009.html>.
- HUANG Xinjie, FANG Sidun, LUO Yingbing, et al. A coordinated voltage management strategy of distribution network considering the state coupling characteristics of energy storage system under bidirectional power flow[J/OL]. Proceedings of the CSEE, 2024: 1-12 [2024-11-24]. <http://kns.cnki.net/kcms/detail/11.2107.tm.20240613.1305.009.html>.
- [22] 廖小兵, 周自强, 乐健, 等. 考虑电池储能老化和需求侧响应的交直流混合配电网动态重构方法[J]. 电力系统保护与控制, 2023, 51(14): 43-54.
- LIAO Xiaobing, ZHOU Ziqiang, LE Jian, et al. Dynamic reconfiguration of an AC/DC hybrid distribution network considering battery energy storage aging and demand response[J]. Power System Protection and Control, 2023, 51(14): 43-54.
- [23] 张凯越, 汤旻安, 杜涵潇, 等. 基于 AMFOA 考虑主观因素的配电网多目标规划[J]. 控制工程, 2021, 28(1): 68-76.
- ZHANG Kaiyue, TANG Min'an, DU Hanxiao, et al. Multi-objective planning of distribution network considering subjective and objective comprehensive factors based on AMFOA[J]. Control Engineering of China, 2021, 28(1): 68-76.
- [24] 班波, 袁至. 基于破圈法与改进布谷鸟算法的两阶段配电网重构[J]. 现代电子技术, 2023, 46(3): 161-166.
- BAN Bo, YUAN Zhi. Two-stage distribution network reconstruction based on breaking circle method and improved cuckoo algorithm[J]. Modern Electronics Technique, 2023, 46(3): 161-166.
- [25] 潘凯岩, 赵瑞锋, 刘尚伟, 等. 考虑电力负荷需求和分布式电源不确定性的配电网重构[J]. 可再生能源, 2024, 42(4): 546-552.
- PAN Kaiyan, ZHAO Ruifeng, LIU Shangwei, et al. Distribution network reconfiguration considering power load demand and distributed power source uncertainty[J]. Renewable Energy Resources, 2024, 42(4): 546-552.
- [26] 王晨, 徐璐辉, 王淑侠, 等. 基于改进大猩猩算法的主动配电网动态重构[J]. 水电能源科学, 2024, 42(1): 211-215.
- WANG Chen, XU Luhui, WANG Shuxia, et al. Dynamic reconfiguration of active distribution network based on improved gorilla troops optimizer[J]. Water Resources and Power, 2024, 42(1): 211-215.
- [27] 江泽昌, 刘天羽, 江秀臣, 等. 考虑 DG 与负荷不确定性的主动配电网重构[J]. 电测与仪表, 2019, 56(19): 76-84.
- JIANG Zechang, LIU Tianyu, JIANG Xiuchen, et al. Reconfiguration of active distribution network considering DG and load uncertainty[J]. Electrical Measurement & Instrumentation, 2019, 56(19): 76-84.
- [28] 吴建旭, 于永进. 基于改进和声搜索算法的多目标配电网重构优化[J]. 电力系统保护与控制, 2021, 49(19): 78-86.

- WU Jianxu, YU Yongjin. Multi-objective distribution network reconfiguration optimization based on an improved harmony search algorithm[J]. *Power System Protection and Control*, 2021, 49(19): 78-86.
- [29] 石立国, 李延真, 胡洋, 等. 提高电动汽车接纳能力的配电网动态重构方法[J]. *供用电*, 2024, 41(4): 89-95.
- SHI Liguó, LI Yanzhen, HU Yang, et al. Dynamic reconfiguration method of distribution network for improving electric vehicle hosting capacity[J]. *Distribution & Utilization*, 2024, 41(4): 89-95.
- [30] 李扬, 韦钢, 马钰, 等. 含电动汽车和分布式电源的主动配电网动态重构[J]. *电力系统自动化*, 2018, 42(5): 102-110.
- LI Yang, WEI Gang, MA Jue, et al. Dynamic reconfiguration of active distribution network considering electric vehicles and distributed generations[J]. *Automation of Electric Power Systems*, 2018, 42(5): 102-110.
- [31] MAHDAVI M, ALHELOU H H, HATZIARGYRIOU N D, et al. Reconfiguration of electric power distribution systems: comprehensive review and classification[J]. *IEEE Access*, 2021, 9: 118502-118527.
- [32] 倪识远, 张林森. 考虑动态重构的主动配电网多目标双层优化调度方法[J]. *电力系统保护与控制*, 2020, 48(20): 38-47.
- NI Shiyuan, ZHANG Linyao. Multi-objective bi-level optimal dispatch method of an active distribution network considering dynamic reconfigurations[J]. *Power System Protection and Control*, 2020, 48(20): 38-47.
- [33] LIU Z F, LIU Y Y, QU G Q, et al. Intra-day dynamic network reconfiguration based on probability analysis considering the deployment of remote control switches[J]. *IEEE Access*, 2019, 7: 145272-145281.
- [34] SARANTAKOS I, GREENWOOD D M, YI J L, et al. A method to include component condition and substation reliability into distribution system reconfiguration[J]. *International Journal of Electrical Power & Energy Systems*, 2019, 109: 122-138.
- [35] 李振坤, 陈星莺, 余昆, 等. 配电网重构的混合粒子群算法[J]. *中国电机工程学报*, 2008, 28(31): 35-41.
- LI Zhenkun, CHEN Xingying, YU Kun, et al. Hybrid particle swarm optimization for distribution network reconfiguration[J]. *Proceedings of the CSEE*, 2008, 28(31): 35-41.
- [36] 马草原, 孙展展, 尹志超, 等. 基于双重混合粒子群算法的配电网重构[J]. *电工技术学报*, 2016, 31(11): 120-128.
- MA Caoyuan, SUN Zhanzhan, YIN Zhichao, et al. Reconfiguration of distribution network based on double hybrid particle swarm algorithm[J]. *Transactions of China Electrotechnical Society*, 2016, 31(11): 120-128.
- [37] 阳晓明, 吕红芳, 朱辉. 基于改进人工鱼群算法的配电网重构[J]. *电测与仪表*, 2020, 57(17): 72-78, 98.
- YANG Xiaoming, LÜ Hongfang, ZHU Hui. Reconfiguration of distribution network based on improved artificial fish swarm algorithm[J]. *Electrical Measurement & Instrumentation*, 2020, 57(17): 72-78, 98.
- [38] WANG B B, ZHU H, XU H H, et al. Distribution network reconfiguration based on NoisyNet deep Q-learning network[J]. *IEEE Access*, 2021, 9: 90358-90365.
- [39] GALLEGO PAREJA L A, LÓPEZ-LEZAMA J M, GÓMEZ CARMONA O. Optimal integration of distribution network reconfiguration and conductor selection in power distribution systems via MILP[J]. *Energies*, 2023, 16(19): 6998.
- [40] MOGHARI P, CHABANLOO R M, TORKAMAN H. Distribution system reconfiguration based on MILP considering voltage stability[J]. *Electric Power Systems Research*, 2023, 222: 109523.
- [41] LEI C, BU S Q, ZHONG J, et al. Distribution network reconfiguration: a disjunctive convex hull approach[J]. *IEEE Transactions on Power Systems*, 2023, 38(6): 5926-5929.
- [42] TAMI Y, SEBAA K, LAHDEB M, et al. Mixed-integer quadratic constrained programming versus quadratic programming methods for distribution network reconfiguration[C]//2019 International Conference on Advanced Electrical Engineering (ICAEE). Algiers, Algeria: IEEE, 2019: 1-5.
- [43] CHEN Y J, CHEN X M, WU Z L, et al. Coordinated optimization of distribution system reconfiguration and reactive power based on mixed integer convex programming[C]//2018 International Conference on Power System Technology (POWERCON). Guangzhou, China: IEEE, 2018: 2015-2020.
- [44] MAHDAVI M, ALHELOU H H, HATZIARGYRIOU N D, et al. An efficient mathematical model for distribution system reconfiguration using AMPL[J]. *IEEE Access*, 2021, 9: 79961-79993.
- [45] YANG T S, GUO Y, DENG L R, et al. A linear branch flow model for radial distribution networks and its application to reactive power optimization and network reconfiguration[J]. *IEEE Transactions on Smart Grid*, 2021, 12(3): 2027-2036.
- [46] MAHDAVI M, ROMERO R. Reconfiguration of radial distribution systems: an efficient mathematical model[J]. *IEEE Latin America Transactions*, 2021, 19(7): 1172-1181.
- [47] MA C X, WANG D, SHAN X, et al. A fault recovery strategy of distribution network based on mixed-integer second-order cone programming[C]//2020 5th Asia Conference on Power and Electrical Engineering (ACPEE). Chengdu, China: IEEE, 2020: 1584-1589.
- [48] 胡维昊, 曹迪, 黄琦, 等. 深度强化学习在配电网优化运行中的应用[J]. *电力系统自动化*, 2023, 47(14): 174-191.
- HU Weihao, CAO Di, HUANG Qi, et al. Application of deep reinforcement learning in optimal operation of distribution network[J]. *Automation of Electric Power Systems*, 2023, 47(14): 174-191.
- [49] TRUONG A V, TON T N, DUONG T L, et al. Reconfigure the distribution network with photovoltaic connection to minimize energy loss based on average branch power and an advanced branch exchange algorithm[J]. *IEEE Access*, 2021, 9: 104572-104581.
- [50] KAZEMI-ROBATI E, SEPASIAN M S. Fast heuristic methods for harmonic minimization using distribution system reconfiguration[J]. *Electric Power Systems Research*, 2020, 181: 106185.
- [51] DEJAMKHOY A, KHANEGHAH Y F, SHAYEGHI H. Modified branch exchange reconfiguration of active distributed network for simultaneous loss reduction and power quality improvement[J]. *International Transactions on Electrical Energy Systems*, 2019, 29(10): e12065.
- [52] SHIRMOHAMMADI D, HONG H W. Reconfiguration of electric distribution networks for resistive line losses reduction[J]. *IEEE Transactions on Power Delivery*, 1989, 4(2): 1492-1498.
- [53] QU Y P, LIU C C, XU J, et al. A global optimum flow pattern for feeder reconfiguration to minimize power losses of unbalanced distribution systems[J]. *International Journal of Electrical Power & Energy Systems*, 2021, 131: 107071.
- [54] ZHAN J P, LIU W J, CHUNG C Y, et al. Switch opening and exchange method for stochastic distribution network reconfiguration[J]. *IEEE Transactions on Smart Grid*, 2020, 11(4): 2995-3007.
- [55] GUO L X, HUANG Z G, WANG J, et al. Distribution network reconfiguration algorithms: a comparative study[C]//2022 IEEE 5th International Conference on Electronics Technology (ICET). Chengdu, China: IEEE, 2022: 472-478.
- [56] FREITAS K B, ARANTES M S, TOLEDO C F M, et al. MIQP model and improvement heuristic for power loss minimization in distribution system with network reconfiguration[J]. *Journal of Heuristics*, 2020,

- 26(1): 59-81.
- [57] LI H W, MAO W J, ZHANG A N, et al. An improved distribution network reconfiguration method based on minimum spanning tree algorithm and heuristic rules[J]. *International Journal of Electrical Power & Energy Systems*, 2016, 82: 466-473.
- [58] JAKUS D, ČAĐENOVIC R, VASILJ J, et al. Optimal reconfiguration of distribution networks using hybrid heuristic-genetic algorithm[J]. *Energies*, 2020, 13(7): 1544.
- [59] JANGDOOST A, KEYPOUR R, GOLMOHAMADI H. Optimization of distribution network reconfiguration by a novel RCA integrated with genetic algorithm[J]. *Energy Systems*, 2021, 12(3): 801-833.
- [60] ESSALLAH S, KHEDHER A. Optimization of distribution system operation by network reconfiguration and DG integration using MPSSO algorithm[J]. *Renewable Energy Focus*, 2020, 34: 37-46.
- [61] HIZARCI H, DEMIREL O, TURKAY B E. Distribution network reconfiguration using time-varying acceleration coefficient assisted binary particle swarm optimization[J]. *Engineering Science and Technology, an International Journal*, 2022, 35: 101230.
- [62] STOJANOVIĆ B, RAJIĆ T. Distribution network reconfiguration and capacitor switching in the presence of wind generators[J]. *Electrical Engineering*, 2022, 104(4): 2249-2266.
- [63] STOJANOVIĆ B, RAJIĆ T, ŠOŠIĆ D. Distribution network reconfiguration and reactive power compensation using a hybrid simulated annealing-minimum spanning tree algorithm[J]. *International Journal of Electrical Power & Energy Systems*, 2023, 147: 108829.
- [64] KOONG G I, MOKHLIS H, JAMIAN J J, et al. Simultaneous network reconfiguration with distributed generation sizing and tap changer adjustment for power loss reduction using imperialist competitive algorithm[J]. *Arabian Journal for Science and Engineering*, 2018, 43(6): 2779-2792.
- [65] PAMBAYUN Y A, ZAINI A, ALI M, et al. Rekonfigurasi jaringan distribusi radial Di Penyulang Purwoasri berbasis modified imperialist competitive algorithms (MICA)[J]. *Jurnal FORTECH*, 2020, 1(2): 74-78.
- [66] DOS SANTOS M V, BRIGATTO G A, GARCÉS L P. Methodology of solution for the distribution network reconfiguration problem based on improved harmony search algorithm[J]. *IET Generation, Transmission & Distribution*, 2020, 14(26): 6526-6533.
- [67] DIAS SANTOS J, MARQUES F, GARCÉS NEGRETE L P, et al. A novel solution method for the distribution network reconfiguration problem based on a search mechanism enhancement of the improved harmony search algorithm[J]. *Energies*, 2022, 15(6): 2083.
- [68] GU S L, XU S L, HAN L, et al. A reconfiguration strategy for active distribution network based on classification and regression tree[C]//2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2). Beijing, China: IEEE, 2018: 1-5.
- [69] SARKAR D, GUNTURI S K. Machine learning enabled steady-state security predictor as deployed for distribution feeder reconfiguration[J]. *Journal of Electrical Engineering & Technology*, 2021, 16(3): 1197-1206.
- [70] DENG W, YANG J Q, TIAN Y, et al. A distribution network reconfiguration method based on the representative daily load curve[C]//2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2). Changsha, China: IEEE, 2019: 2811-2815.
- [71] YANG Q, ZHAO E G, MA X, et al. Optimal reconfiguration of distribution network based on deep fuzzy C-means clustering algorithm[J]. *Journal of Physics: Conference Series*, 2023, 2497: 012005.
- [72] MAHDAVI M, JAVADI M, WANG F, et al. An accurate evaluation of consumption pattern in reconfiguration of electrical energy distribution systems[C]//2021 IEEE Industry Applications Society Annual Meeting (IAS). Vancouver, Canada: IEEE, 2021: 1-7.
- [73] YOUSSEF H H, MOKHLIS H B, TALIP M S A, et al. Distribution network reconfiguration based on artificial network reconfiguration for variable load profile[J]. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2020, 28(5): 3013-3035.
- [74] HUANG W J, ZHENG W Y, HILL D J. Distribution network reconfiguration for short-term voltage stability enhancement: an efficient deep learning approach[J]. *IEEE Transactions on Smart Grid*, 2021, 12(6): 5385-5395.
- [75] HUANG W J, ZHAO C H. Deep-learning-aided voltage-stability-enhancing stochastic distribution network reconfiguration[J]. *IEEE Transactions on Power Systems*, 2024, 39(2): 2827-2836.
- [76] ZHENG W Y, HUANG W J, HILL D J, et al. An adaptive distributionally robust model for three-phase distribution network reconfiguration[J]. *IEEE Transactions on Smart Grid*, 2021, 12(2): 1224-1237.
- [77] LU C, LIU Y T, TIAN Y, et al. Research on distribution network topology reconfiguration based on LSTM neural network[C]//2021 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA). Dalian, China: IEEE, 2021: 9-13.
- [78] YUE D, HE Z W, DOU C X. Cloud-edge collaboration-based distribution network reconfiguration for voltage preventive control[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(12): 11542-11552.
- [79] 蔡田田, 姚浩, 杨英杰, 等. 基于云-边协同的配电网快速供电恢复智能决策方法[J]. *电力系统保护与控制*, 2023, 51(19): 94-103. CAI Tiantian, YAO Hao, YANG Yingjie, et al. Cloud-edge collaboration-based supply restoration intelligent decision-making method[J]. *Power System Protection and Control*, 2023, 51(19): 94-103.
- [80] GAO Y Q, SHI J, WANG W, et al. Dynamic distribution network reconfiguration using reinforcement learning[C]//2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (Smart Grid Comm). Beijing, China: IEEE, 2019: 1-7.
- [81] KUNDAČINA O B, VIDOVIĆ P M, PETKOVIĆ M R. Solving dynamic distribution network reconfiguration using deep reinforcement learning[J]. *Electrical Engineering*, 2022, 104(3): 1487-1501.
- [82] GAO Y Q, WANG W, SHI J, et al. Batch-constrained reinforcement learning for dynamic distribution network reconfiguration[J]. *IEEE Transactions on Smart Grid*, 2020, 11(6): 5357-5369.
- [83] ZHAO B C, HAN X S, MA Y R, et al. Fast reconfiguration of distribution network based on deep reinforcement learning algorithm[J]. *IOP Conference Series: Earth and Environmental Science*, 2020, 571: 012023.
- [84] LIN Z, JIANG C X, LU Y J, et al. Active distribution network reconfiguration with renewable energy based on multi-agent deep reinforcement learning[C]//2023 6th International Conference on Energy, Electrical and Power Engineering (CEEPE). Guangzhou, China: IEEE, 2023: 535-542.
- [85] XU J, GAO H J, WANG R J, et al. Real-time operation optimization in active distribution networks based on multi-agent deep reinforcement learning[J]. *Journal of Modern Power Systems and Clean Energy*, 2024, 12(3): 886-899.
- [86] GHOLIZADEH N, KAZEMI N, MUSILEK P. A comparative study of reinforcement learning algorithms for distribution network reconfiguration with deep Q-learning-based action sampling[J]. *IEEE Access*, 2023, 11: 13714-13723.
- [87] 董雷, 吴怡, 张涛, 等. 基于强化学习的含智能软开关主动配电网双层优化方法[J]. *电力系统自动化*, 2023, 47(6): 59-68. DONG Lei, WU Yi, ZHANG Tao, et al. Reinforcement learning based

- double-layer optimization method for active distribution network with soft open point[J]. Automation of Electric Power Systems, 2023, 47(6): 59-68.
- [88] YIN Z Y, WANG S X, ZHAO Q Y. Sequential reconfiguration of unbalanced distribution network with soft open points based on deep reinforcement learning[J]. Journal of Modern Power Systems and Clean Energy, 2023, 11(1): 107-119.
- [89] JACOB R A, PAUL S, LI W Y, et al. Reconfiguring unbalanced distribution networks using reinforcement learning over graphs[C]//2022 IEEE Texas Power and Energy Conference (TPEC). College Station, USA: IEEE, 2022: 1-6.
- [90] LI Y Z, HAO G K, LIU Y, et al. Many-objective distribution network reconfiguration via deep reinforcement learning assisted optimization algorithm[J]. IEEE Transactions on Power Delivery, 2022, 37(3): 2230-2244.
- [91] GAUTAM M, BHUSAL N, BENIDRIS M. Deep Q-Learning-based distribution network reconfiguration for reliability improvement[C]//2022 IEEE/PES Transmission and Distribution Conference and Exposition (T&D). New Orleans, USA: IEEE, 2022: 1-5.
- [92] GAUTAM M, BENIDRIS M. Distribution network reconfiguration using deep reinforcement learning[C]//2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). Manchester, UK: IEEE, 2022: 1-6.
- [93] MALEKSHAH S, RASOULI A, MALEKSHAH Y, et al. Reliability-driven distribution power network dynamic reconfiguration in presence of distributed generation by the deep reinforcement learning method[J]. Alexandria Engineering Journal, 2022, 61(8): 6541-6556.
- [94] BUI V H, SU W C. Real-time operation of distribution network: a deep reinforcement learning-based reconfiguration approach[J]. Sustainable Energy Technologies and Assessments, 2022, 50: 101841.
- [95] 徐何军. 基于神经网络的配电网故障诊断及重构优化研究[D]. 武汉: 华中科技大学, 2020.
- XU Hejun. Research on fault diagnosis and reconfiguration optimization of distribution network based on graph neural network[D]. Wuhan, China: Huazhong University of Science and Technology, 2020.
- [96] 闫冬, 彭国政, 高海龙, 等. 基于深度强化学习组合优化的配电网拓扑控制研究[J]. 电网技术, 2022, 46(7): 2547-2554.
- YAN Dong, PENG Guozheng, GAO Hailong, et al. Research on distribution network topology control based on deep reinforcement learning combinatorial optimization[J]. Power System Technology, 2022, 46(7): 2547-2554.
- [97] 张沛, 陈玉鑫, 王光华, 等. 基于图强化学习的配电网故障恢复决策[J]. 电力系统自动化, 2024, 48(2): 151-158.
- ZHANG Pei, CHEN Yuxin, WANG Guanghua, et al. Fault recovery decision of distribution network based on graph reinforcement learning[J]. Automation of Electric Power Systems, 2024, 48(2): 151-158.



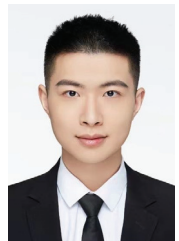
江昌旭

1990—, 男, 博士, 副教授, 硕导  
主要从事电力系统运行优化、强化学习(人工智能)  
以及电力交通协同优化方面的研究  
E-mail: cxjiang@fzu.edu.cn;

JIANG Changxu

Ph.D.

Associate professor



郭辰

2000—, 男, 硕士  
研究方向为强化学习、配电网重构  
E-mail: 3496028218@qq.com

GUO Chen

刘晨曦

2000—, 男, 硕士  
研究方向为人工智能技术在电力系统中的应用  
E-mail: 1584891792@qq.com



LIU Chenxi



林俊杰(通信作者)

1992—, 男, 博士, 副教授, 硕导  
主要从事电力系统分析与控制技术、状态估计等  
方面的研究  
E-mail: linjunjie@fzu.edu.cn

LIN Junjie

Ph.D.

Associate professor  
Corresponding author

邵振国

1970—, 男, 博士, 教授, 博导  
主要从事综合能源系统协同控制方法、电力大数据  
分析与挖掘方面的研究  
E-mail: shao.zg@fzu.edu.cn

SHAO Zhenguo

Ph.D., Professor

附录 A

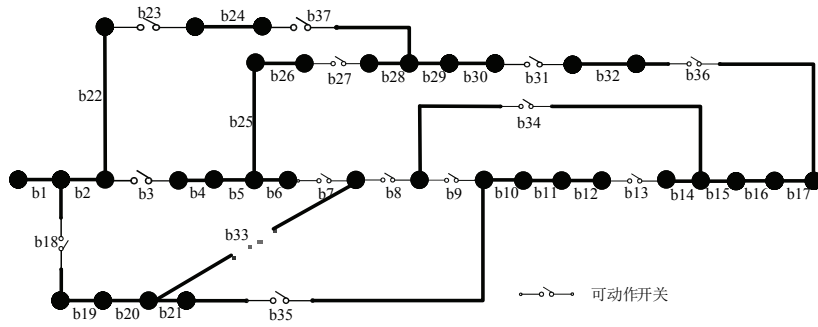
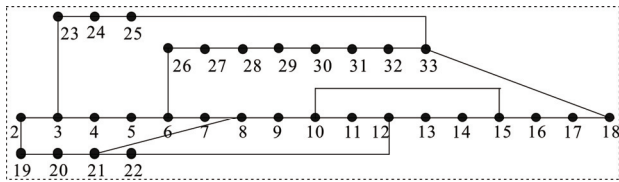
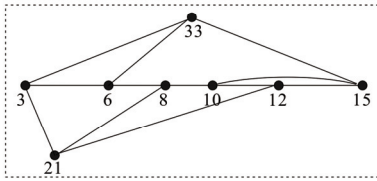


图 A1 IEEE 33 节点系统

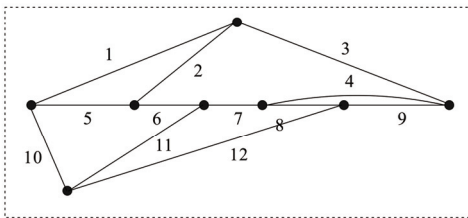
Fig.A1 IEEE 33 node system



(a) IEEE 33节点系统简化图



(b) IEEE 33节点系统同胚图



(c) IEEE 33节点系统基于支路组的编码图

图 A2 基于支路组的编码方式图

Fig.A2 Encoding method diagram based on branch groups

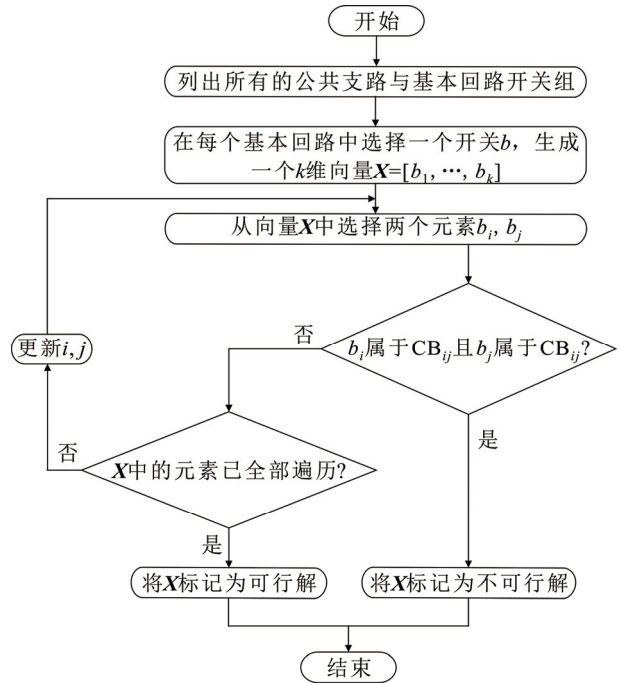


图 A3 基于基本回路的编码方式流程图

Fig.A3 Flow chart of coding method based on basic circuit

表 A1 IEEE 33 节点标准测试系统十进制开关编码表

Table A1 IEEE 33-bus standard test system decimal switch

encoding

基本回路	可动作开关所在支路	十进制编码
F <sub>1</sub>	b3, b23, b27, b37	1~4
F <sub>2</sub>	b7, b8, b27, b31, b34, b36	1~6
F <sub>3</sub>	b9, b13, b34	1~3
F <sub>4</sub>	b3, b7, b18, b33	1~4
F <sub>5</sub>	b8, b9, b33, b35	1~4

表 A2 基于各强化学习算法的主动配电网动态重构策略分析

Table A2 Analysis of dynamic reconfiguration strategies for active distribution networks based on various reinforce learning algorithms

强化学习算法类别	具体算法	优点	缺点	改进方向
传统强化学习算法	Q 学习	无需准确的配电网参数进行模型构建	使用表格存储 $Q$ 值在面对大规模配电网会出现维数灾难问题	使用函数对 $Q$ 函数进行拟合
	近似动态规划算法	通过函数拟合可避免在面对大规模配电网时维数灾难问题	简单的多项式逼近误差较大	使用神经网络对 $Q$ 函数进行拟合
基于价值的强化学习算法	DQN	深度神经网络拟合函数的误差较小	容易出现过高估计风险	使用目标网络估计目标 $Q$ 值
	Double DQN	过高估计的概率比 DQN 低	维护两个神经网络所需的计算复杂度高	收敛速度慢
	竞争 DQN	计算复杂度低, 动作输出精确度更高	仍有过高估计的风险	加入优先经验回放机制
基于策略的强化学习算法	TRPO	收敛速度快, 处理连续动作空间效果较优	稳定性能差, 需要的实际训练样本数量较大	与基于价值的强化学习下算法进行结合
	PPO			
价值与策略结合的强化学习算法	AC	收敛速度快, 稳定性好	容易出现过高估计风险	使用目标网络估计目标 $Q$ 值
	DDPG	过高估计的概率比 AC 低	收敛速度快且稳定性好	确定性动作选择策略
	TD3	减轻 $Q$ 值过高估计问题	相较于 AC 与 DDPG, 稳定性不足	对 $Q$ 函数进行优化及延迟策略更新
多智能体强化学习算法	MADDPG	通过将每一个回路或开关视为一个智能体可以提高算法效率	智能体的竞争与冲突会影响算法性能	设置合适的奖励函数与协作机制
	MATD3			

## 附录 B

ADNDR 数学模型中包含了负荷的不确定性与 DG 出力的不确定性。因此, 该数学模型的本质是一个高维、非线性的随机优化数学模型。传统数学算法在解决该问题的过程中存在着计算速度慢、无法有效处理大规模随机变量等问题, 启发式算法则容易陷入局部最优解。为了更加快速、准确地对 ADNDR 数学模型进行求解, 现有研究引入了强化学习算法。该算法通过与环境(配电网)不断地进行交互, 观察环境状态(节点有功功率、节点无功功率以及节点电压等), 从环境反馈的奖惩信号中进行自适应学习, 以此获得最大的长期累计平均回报和相应的行为策略(断开或闭合某些开关)。在强化学习中, ADNDR 问题通常被建模为一个马尔可夫决策过程(markov decision process, MDP)。从形式上看, MDP 由五元组  $(S, A, P, R, \gamma)$  组成,  $S$  是状态空间,  $A$  是智能体可选择的动作空间,  $P$  是满足马尔可夫性质的转移概率,  $R$  为奖励函数,  $\gamma \in (0, 1)$  为折扣因子。在 ADNDR 策略研究中, 强化学习算法的组成部分如表 B1 所示。

表 B1 ADNDR 策略研究中的强化学习算法组成部分

Table B1 Reinforce learning algorithm components in

ADNDR strategy research

RL 算法组成部分	ADNDR 策略研究中的对应部分
智能体	配电网操作员
环境	配电网
状态 $s \in S$	网络损耗、电压偏移以及运行费用等
动作 $a \in A$	断开或闭合配电网中的某些开关
奖励 $r$	与数学模型中的目标函数类似

强化学习是一种免模型的算法, 具有较强的学习能力、适应性和可扩展性, 能够有效地在动态环境中学习到最优的策略, 因此能够有效地解决主动配电网复杂环境下的动态重构问题。因此, 许多学者基于强化学习算法对 ADNDR 开展了深入研究。文献[B1]提出了一种基于 Q 学习的 ADNDR 策略, 将 ADNDR 构建为 MDP, 并将每一步的  $Q$  值存储在  $Q$  表格中。但在高维的优化问题中, 状态与动作的维度极高, 使用表格对  $Q$  值进行存储会出现“维数灾难”问题。文献[B2]使用近似动态规划算法拟合价值函数对 ADNDR 进行求解, 避免了维数灾难问题。但近似动态规划算法通常使用一些简单的函数逼近方法, 这些方法的表达能力有限, 可能无法很

好地拟合复杂的价值函数。

与上述方法相比, DRL 算法将强化学习算法与深度学习相结合, 通过 DNN 近似价值函数或策略函数, 从而更准确地估计状态和动作的价值, 同时, DRL 算法通过不断与配电网进行交互, 获取当前时刻的配电网状态与奖励函数, 并据此优化重构策略。这种方式能够快速适应电网状态的变化以及负荷与 DG 出力的不确定性, 因而有着较高的求解精度。并且, DRL 算法避免了传统算法需要针对每个实例从头迭代计算的局限, 通过探索与利用的平衡实现高效学习, 利用已有的经验学习不同时刻配电网的共性特征, 同时探索新的可能性。并将训练后的模型直接用于求解新的重构策略, 使得智能体能够快速求解出最优策略, 大大提高了求解效率。此外, DRL 算法通过与配电网的交互动态地更新重构策略, 不依赖于配电网的具体规模和结构, 因此具有极强的泛化性能。并且, DRL 算法可以在探索新动作以发现更好的策略和利用已知信息以最大化长期收益之间找到平衡, 这种平衡使得 DRL 算法有着较强的全局搜索能力。因此, DRL 算法在求解 ADNDR 策略时能够兼顾不确定因素适应性、求解精度、求解效率、泛化性能与全局搜索能力。

基于 DRL 算法的 ADNDR 基本框架如图 B1 所示。从图中可以得知, 根据智能体动作选取的方式, 可以将常规的 DRL 算法分为基于价值的 DRL 算法、基于策略的 DRL 算法以及结合价值与策略的 DRL 算法 3 类。

#### 附录 B1 基于价值的 DRL 算法

基于价值的 DRL 算法评估每个状态  $s$  或状态动作对  $(s, a)$  的价值, 以此指导智能体的决策过程, 使其选择能够最大化未来累积奖励的动作。动作价值函数  $Q^\pi(s, a)$  是状态-动作对  $(s, a)$  到动作价值  $Q$  的映射, 表示从状态  $s$  开始, 采取动作  $a$  并遵循策略  $\pi$  所获得的长期回报预期值。动作价值函数直接给出了在给定状态下执行特定动作的预期累积奖励, 这使得智能体能够直接根据  $Q$  值选择最优动作, 而无需通过额外的计算或推断。常见的基于价值的方法包括 Q 学习算法、深度 Q 网络(deep Q network, DQN) 算法以及其改进算法, 如双深度 Q 网络算法(double deep Q network, DDQN)等。

由于动作价值函数十分平滑, 对状态的微小变化不敏感, 因此, 基于价值的方法能够更加稳定地应对配电网状态的微小波动, 避免频繁的策略切换,

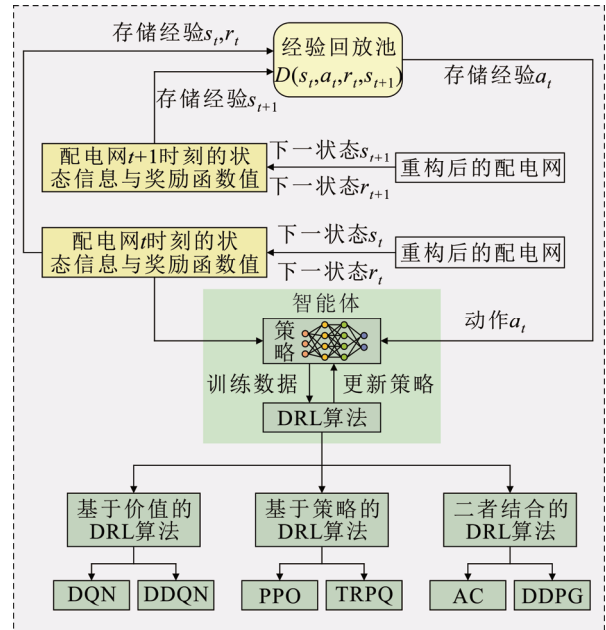


Fig.B1 Framework of active distribution network dynamic reconfiguration algorithm based on deep reinforcement learning

从而能提供更加稳定的重构策略。此外, 基于价值的算法通常具有较高的样本效率, 因为它们可以利用动作价值函数来评估不同动作的好坏, 从而更有效地选择有利动作。这意味着算法能够更快地学习到有效的重构策略, 减少训练时间和计算资源消耗。并且, 基于价值的算法的可解释性较强, 因为动作价值函数直接反映了状态或状态-动作对的好坏。这有助于帮助决策者需要理解算法为什么选择某个重构策略, 以便进行进一步的优化和调整。

#### 附录 B2 基于策略的 DRL 算法

在基于价值的 DRL 算法中, 动作通常通过最大化估计的  $Q$  值进行选择, 这会导致某些动作的  $Q$  值被高估, 从而使得动作价值函数出现过高估计的问题。基于策略的 DRL 算法直接学习最优策略, 通过参数化策略函数获得最大的长期回报, 避免了过度估计问题。常见的基于策略的方法包括信任区域策略优化(trust region policy optimization, TRPO)算法与近端策略优化(proximal policy optimization, PPO)算法等。

基于策略的深度强化学习算法, 能够直接学习并输出动作的概率分布, 因此非常适合处理具有连续动作空间的问题。在求解主动配电网动态重构与无功补偿协同优化策略时, 基于策略的方法可用于求解连续的无功补偿设备(静止无功补偿器、DG 逆变器)的无功补偿量。此外, 在处理大规模配电网的

重构策略时，由于基于策略的方法不需要为每个可能的动作计算价值，因此其通常比基于价值的方法更为高效。

### 附录 B3 价值和策略结合的 DRL 算法

在基于价值的 DRL 算法中，价值函数的估计和策略改进是分开的。价值函数的准确性直接影响策略的质量，而价值函数的学习往往需要较长时间，因此其收敛速度较慢。基于策略的 DRL 算法在策略梯度的估计过程中会受到环境噪声的影响，因此往往具有较高的方差，这会导致训练过程不稳定。价值和策略结合的 DRL 算法利用价值函数指导策略学习，并通过策略学习更新价值函数。这种相互补充的方式使得算法能够在提高收敛速度的同时，减少策略梯度的方差，使得训练过程更加稳定。常见的价值和策略结合的方法有演员评论家(actor critic, AC)算法与深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法等。其中策略函数被称为“演员”，价值函数被称为“评论家”。

在求解 ADNDR 的过程中，评论家网络通过动作价值函数评估不同重构动作对配电网性能的影响，为演员网络提供反馈信号，从而减少策略估计的方差，使得演员的策略更新更为稳定。演员网络可以根据当前配电网的状态与评论家网络的反馈输出最优重构动作，以满足优化目标。这种策略与动作价值函数同时学习的方式，能够提高算法的求解效率。

### 附录 B4 图强化学习算法

传统的 DNN 通常将输入数据视为独立的向量或矩阵，而对于配电网这类具有明显图结构特征的环境，这种方式难以直接表达状态与动作之间的复杂图结构关系。图强化学习算法(graph reinforce learning, GRL)使用图神经网络(graph neural network, GNN)与强化学习算法进行结合，使智能体可以在具有图结构数据的环境中做出决策。GRL 首先将环境建模为一个图  $G(V, E)$ ，图中的节点  $V$  可以表示环境中的实体或对象，而图中的边  $E$  则表示这些实体或对象之间的关系。之后，使用 GNN 提取图中的节点和边的特征并将其转化为高维的向量表示，这些向量可以被视为强化学习中的状态  $s$ 。进而，使用上述 DRL 算法指导智能体的决策过程，智能体通过与环境的交互收集经验，并使用这些经验更新其策略或价值函数。

常见 GRL 算法有图卷积网络深度强化学习(graph convolutional network deep reinforce learning,

GCNDRL)算法与图注意力网络深度强化学习(graph attention networks deep reinforce learning, GATDRL)算法。GCNDRL 在处理图结构数据时，对所有的邻居节点一视同仁，没有考虑节点之间的重要性差异。GATDRL 使用图注意力网络(graph attention networks, GAT)拟合动作价值函数。GAT 的信息传播规则主要基于注意力机制，并将其用于计算图中节点之间的注意力系数，并根据这些系数聚合邻居节点的信息<sup>[B3]</sup>。主动配电网在重构过程中节点间的连接方式将频繁发生变化，GATDRL 通过注意力机制可以动态地调整节点间的注意力系数，从而更好的适应这种变化。

### 附录 B5 多智能体深度强化学习算法

在求解 ADNDR 这类高维的优化问题时，动作空间的维度会随着待求解问题维度的增加而增加，使用传统的单智能体 DRL 算法对其进行求解将降低其求解效率。多智能体深度强化学习(multi-agent deep reinforce learning, MADRL)算法将复杂的高维优化问题分解为多个子任务，每个子任务由不同的智能体负责处理<sup>[B4]</sup>。例如为每个开关或每个基本回路配置一个智能体，用于求解局部的重构策略，均可以减少每个智能体需要处理的动作空间维度。此外，MADRL 中的智能体通过协同学习来共同解决优化问题，它们可以共享彼此的经验、策略和状态，从而加速学习过程。并且，MADRL 算法可以并行求解每个智能体的最优动作，大大提升了求解效率，这在大规模配电网中尤为重要。MADRL 算法框架如图 B2 所示，常见的 MADRL 算法有多智能体深度确定策略梯度(multi agent deep deterministic policy gradient, MADDPG)算法与多智能体双延迟深度确定性策略梯度(multi agent twin delayed deep deterministic policy gradient, MATD3)算法。

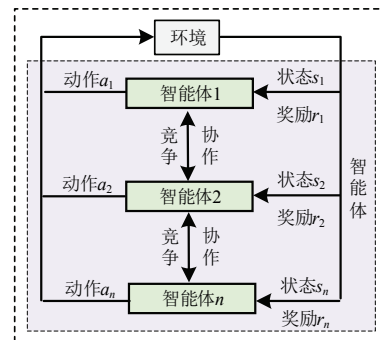


图 B2 多智能体强化学习算法框架图

Fig.B2 Framework diagram of multi-agent reinforcement learning algorithm

- [B1] VLACHOGIANNIS J G, HATZIARGYRIOU N. Reinforcement learning (RL) to optimal reconfiguration of radial distribution system (RDS)[C]//Third Hellenic Conference on AI. Samos, Greece: Springer, 2004: 439-446.
- [B2] WANG C, LEI S B, JU P, et al. MDP-based distribution network reconfiguration with renewable distributed generation: approximate dynamic programming approach[J]. IEEE Transactions on Smart Grid, 2020, 11(4): 3620-3631.
- [B3] HUANG P, GUO J W, LIU S, et al. Explainable train delay propagation: a graph attention network approach[J]. Transportation Research Part E: Logistics and Transportation Review, 2024, 184: 103457.
- [B4] 江昌旭, 卢玥君, 邵振国, 等. 基于图神经网络多智能体强化学习的电力-交通融合网协同优化运行[J]. 高电压技术, 2023, 49(11): 4622-4631.
- JIANG Changxu, LU Yuejun, SHAO Zhenguo, et al. Collaborative optimization operation of integrated electric power and traffic network based on graph neural network multi-agent reinforcement learning[J]. High Voltage Engineering, 2023, 49(11): 4622-4631.