

基于数据驱动与物理模型的主动配电网双时间尺度协调优化

张 剑¹, 崔明建², 姚潇毅³, 何怡刚⁴

(1. 合肥工业大学电气与自动化工程学院, 安徽省合肥市 230009;

2. 天津大学电气自动化与信息工程学院, 天津市 300072;

3. 国网安徽省电力有限公司蚌埠供电公司, 安徽省蚌埠市 233000;

4. 武汉大学电气与自动化学院, 湖北省武汉市 430072)

摘要: 高比例间歇性分布式电源与电动汽车接入配电网时,容易导致功率与电压频繁、快速、剧烈波动。文中结合数据驱动与物理建模方法,提出了一种配电网双时间尺度有功无功协调优化策略。针对短时间尺度(分钟级或秒级)的功率波动,以静止无功补偿器、分布式电源无功功率为决策变量,以网损最小为目标函数,计及物理约束,针对平衡与不平衡配电网分别构建了二阶锥与二次规划模型。针对长时间尺度(小时级)的优化,以有载调压变压器分接头变比、可投切电容电抗器挡位、储能系统充放电功率为动作,以网损为代价,计及节点电压越限惩罚,构建了马尔可夫决策过程。为克服连续-离散动作空间维数灾,采用一种基于松弛-预报-校正的深度确定性策略梯度强化学习求解算法。通过 IEEE 33 节点与 IEEE 123 节点配电系统验证了所提方法的有效性。

关键词: 主动配电网; 分布式电源; 深度强化学习; 二阶锥规划; 二次规划

0 引言

计及源荷不确定性,传统基于物理模型的配电网多时段有功无功协调优化属于大规模混合整数非凸非线性随机或鲁棒优化,求解复杂度随配电网拓扑规模与可调设备数量增加呈指数增长,属于非确定性多项式(non-deterministic polynomial, NP)难题。同时,分布式电源(distributed generator, DG)逆变器、有载调压变压器(on-load tap changer, OLTC)分接头、可投切电容电抗器(switchable capacitor reactor, SCR)、储能系统(energy storage system, ESS)、静止无功补偿器(static var compensator, SVC)等可调设备动作速度与调控方式差异很大,使得配电网有功无功协调优化面临维数高、建模困难、求解慢等难题^[1-3]。

基于数据驱动的方法不依赖于精确的配电网模型,易于处理非凸非线性约束、离散变量与源荷不确定性,能够保证解的(近似)最优性,计算速度满足大规模配电网在线实时优化快速需求^[4-5]。深度强化学习(deep reinforcement learning, DRL)基于深度学习强大的感知能力提取复杂、高维环境特征,结合

强化学习与环境交互,完成序贯决策过程,在改进学习性能方面表现出优越性。近年来,基于DRL的配电网无功优化受到了众多专家与学者的关注^[6-8]。

文献[9]利用Q表格强化学习算法在满足操作约束的同时学习一组控制动作进行无功优化,但该算法在处理具有大量状态或动作的任务时效率不高。文献[10]采用批量强化学习设定OLTC变比调节电压,但这种方法需要人工设计特征。文献[11]采用基于蒙特卡洛搜索树的强化学习算法协调调度ESS充放电功率,解决大量光伏接入配电网导致的过电压问题。上述3种方法均未利用神经网络强大的感知与函数逼近能力。文献[12]采用深度Q学习网络(deep Q-learning network, DQN)算法优化电容器的投切,但DQN算法只能处理离散动作控制问题。文献[13]提出了一种OLTC、电容器无功功率-电压优化的安全异轨DRL算法,但未利用DG逆变器参与调节。文献[14]采用多智能体深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法协调控制光伏逆变器无功调节电压。

离散与连续可调设备动作速度不同,适用于不同时间尺度电压问题,如何在不确定性源荷下协调是难点问题。文献[15]构建了离散与连续可调设备长、短时间尺度在线运行方案。长时间尺度马尔可夫决策过程(Markov decision process, MDP)采用

收稿日期: 2022-12-15; 修回日期: 2023-04-27。

上网日期: 2023-08-25。

国家自然科学基金资助项目(52207130)。

DQN算法求解,短时间尺度MDP采用优势执行器-评价器(advantage actor-critic, A2C)算法求解。目前,DQN算法广泛应用于配电网长时间尺度无功优化。然而,当离散可调设备数量较多时,DQN算法存在离散动作空间维数灾。针对此问题,文献[16]构建了不平衡配电网无功优化多智能体MDP,采用多智能体DQN算法求解。尽管多智能体DRL能够克服离散动作空间维数灾,但多智能体训练过程收敛速度远慢于单智能体^[17-18]。该方法将连续动作变量,即DG无功功率离散化,不同类型可调设备动作时间间隔设定为相同,极大增加了离散动作数量,降低了灵活性、最优性与经济效益,难以有效处理短时间尺度电压越限问题。

本文提出了一种配电网双时间尺度有功无功协调优化策略。针对单智能体连续-离散动作空间维数灾问题,采用一种改进DDPG算法求解。在给定长时间尺度MDP离散可调设备与ESS动作值后,求解长时间尺度(小时级)内每个短时间尺度(分钟级或秒级)凸优化物理模型得到的最优目标函数值累加后作为MDP的代价。因此,数据驱动与物理建模方法融为一体,保证解的(近似)最优性。

1 双时间尺度有功与无功协调优化的模型

1.1 系统模型

如附录A图A1所示,将每天划分为 N_{τ} 个时段,记为 $\tau=1, 2, \dots, N_{\tau}$ 。将每个时段 τ 再细分为 N_{τ} 个时隙,记为 $t=1, 2, \dots, N_{\tau}$ 。每个时段 τ 的持续时间为1h。每个时隙 t 的持续时间为几分钟或几秒钟。为应对长时间尺度源荷功率缓慢变化导致的网损增加、电压越限问题,OLTC变比、SCR挡位、ESS充放电功率设定在时段 $\tau-1$ 的末尾至时段 τ 的开始前完成调整,此后保持不变,直到时段 τ 末尾再重新调整。本文假定OLTC变比、SCR挡位、ESS充放电功率在时段 τ 内的每个时隙 t 的值相同,但是从时段 $\tau-1$ 至 τ 的值可变。本文将ESS充放电功率设置为长时间尺度动作变量是因为,若设置为短时间尺度优化变量将会造成处于不同时隙 t 的优化变量耦合,使得基于物理模型的优化问题规模急剧增加,求解速度难以满足实时控制需求。

为应对高比例风电、光伏与快充电动汽车(electric vehicle, EV)接入配电网导致的短时间尺度功率、电压频繁、快速、剧烈波动问题,SVC、DG逆变器无功功率设定在每个时隙 t 的起始时刻进行调节。此外,为降低模型复杂度,本文假定节点 k 的有功功率 $p_k^d(\tau, t)$ 、无功功率 $q_k^d(\tau, t)$ 与DG有功出力 $p_k^g(\tau, t)$ 在每个时隙 t 内恒定,但从时隙 t 到 $t+1$

可变。

1.2 双时间尺度有功无功协调优化

不同类型可调设备动作速度差异很大,针对此问题,本文提出的双时间尺度有功无功协调优化控制策略需要求解以下随机优化问题:

$$\max_{\varphi(\tau, t), \phi(\tau)} E\left(\sum_{\tau=1}^{\infty} \gamma^{\tau} r(\tau)\right) \quad (1)$$

式中: $E(\cdot)$ 为数学期望函数; $\gamma \in (0, 1)$ 为折扣因子; $\phi(\tau), \varphi(\tau, t)$ 分别为长、短时间尺度优化变量; $r(\tau)$ 为奖励。

目标函数数学期望最大值取决于离散决策变量SCR挡位 $y_k^{\text{SCR}}(\tau)$ 、OLTC变比 $n_{lk}^{\text{OLTC}}(\tau)$ 、连续决策变量ESS充放电功率 $p_k^{\text{ESS}}(\tau)$ 在时段 τ 的优化值,以及连续决策变量SVC的无功功率 $q_k^{\text{svc}}(\tau, t)$ 、DG无功功率 $q_k^d(\tau, t)$ 在时隙 (τ, t) 的优化设定值。离散决策变量向量SCR挡位 $\mathbf{y}^{\text{SCR}}(\tau)$ 、OLTC变比 $\mathbf{n}^{\text{OLTC}}(\tau)$ 使得式(1)是非凸的,求解十分困难,属于NP难题。由于必须计及高比例间歇性DG与快充EV实时功率的强不确定性与时变性,各时段、时隙决策变量不是一次性确定的。因此,式(1)不属于日前规划,而属于多时段、多时隙实时序贯决策与控制模型,求解式(1)面临巨大挑战。

针对上述问题,本文设计了一种算法,通过不断观察当前时隙 t 的负荷有功功率向量 $\mathbf{p}^d(\tau, t)$ 、无功功率向量 $\mathbf{q}^d(\tau, t)$ 与DG有功功率向量 $\mathbf{p}^g(\tau, t)$,得到式(1)的(近似)最优值。尽管上述功率随机过程分布函数未知,但其在当前时隙 t 的具体实现值可精确预测^[19]。因此,本文将结合配电网物理模型与数据驱动方法实时求解式(1)。具体而言,在时段 $\tau-1$ 的末尾,OLTC变比、SCR挡位与ESS充放电功率长时间尺度(近似)最优设定值采用DRL算法根据时段 $\tau-1$ 的配电网状态数据学习得到;在时段 τ 内每个时隙 t 的起始时刻,SVC、DG逆变器无功最优设定值在给定OLTC变比、SCR挡位与ESS充放电功率的情况下,通过构建与求解单一时隙凸规划物理模型给出。时段 τ 内各个时隙 t 的物理模型目标函数优化结果累加值为该时段 τ 的MDP代价。短、长时间尺度模型的具体构建与协调、融合方法见附录B与第2章。

2 长时间尺度ESS与离散可调设备设置

长时间尺度有功无功协调优化模型以OLTC变比、SCR挡位、ESS充放电功率为动作变量。现有方法主要基于启发式、半定规划、二阶锥规划(second-order cone programming, SOCP)凸松弛技术构建多时段混合整数非凸非线性模型。求解结果

不能保证最优性,计算复杂度高,要求计算机具有较大内存空间。本文基于人工智能领域的最新进展,采用基于数据驱动的方法构建MDP,从负荷与间歇性DG功率未知动态分布中求出(近似)最优解,克服了传统优化方法计算复杂度高的弊端,能够满足在线实时控制要求。

2.1 基于数据驱动的建模方法

由附录B式(B3)、式(B4)、式(B6)、式(B17)、式(B18)、式(B28)与附录A图A1可以看出,时段 $\tau-1$ 末尾的OLTC变比、SCR挡位、ESS充放电功率设定值(长时间尺度学习)对时段 τ 内每个时隙 t 的起始时刻SVC、DG逆变器无功设定值(短时间尺度凸优化)均具有重大影响。反之,时段 τ 内每个时隙 t 的起始时刻SVC、DG逆变器无功设定值经由奖励对后续时段OLTC变比、SCR挡位、ESS充放电功率设定值产生影响。这种双向作用十分适合采用强化学习方法求解。

将负荷与间歇性DG功率动态特性建模为MDP,OLTC变比、SCR挡位、ESS充放电功率(近似)最优设定值可采用强化学习算法高效求解。MDP定义为一个6元组 (S, A, R, P, π, J) 。其中, S 为状态集合; A 为动作集合; P 为转移矩阵集合,即根据当前状态 $s \in S$ 与动作 $a \in A$ 给出下一个状态出现的概率; R 为奖励集合,实数 $r \in R$ 为奖励,使得对于任意 $s \in S$ 与 $a \in A, r = r(s, a)|_{s \in S, a \in A}$ 为强化学习智能体在状态 s 采取动作 a 后的瞬时奖励; J 为长期回报集合,时段 τ 的回报 $J_\tau = E(\sum \gamma^r r(\tau))$; π 为策略,为强化学习智能体在状态采取不同动作的概率。

1)动作集合 A :每一种可能的OLTC变比、SCR挡位与ESS充放电功率组合构成一个动作。动作的每个分量对应于一个可调设备。动作的数量随着OLTC、SCR与ESS数量增加呈指数增长。时段 τ 的动作 a_τ 定义为:

$$a_\tau = [(\mathbf{n}^{\text{OLTC}}(\tau))^T, (\mathbf{y}^{\text{SCR}}(\tau))^T, (\mathbf{p}^{\text{ESS}}(\tau))^T]^T \quad (2)$$

式中: $\mathbf{p}^{\text{ESS}}(\tau)$ 为ESS的功率向量。

2)状态集合 S :状态包括时段 $\tau-1$ 的OLTC变比、SCR挡位、ESS功率 $\mathbf{p}^{\text{ESS}}(\tau-1)$ 、每个节点吸收的平均功率向量 $\bar{\mathbf{p}}(\tau-1) + j\bar{\mathbf{q}}(\tau-1)$ 与时段 $\tau-1$ 的末尾ESS存储的电量向量 $\mathbf{E}^{\text{ESS}}(\tau-1)$ 。其中, $\bar{\mathbf{p}}(\tau-1)$ 为在时段 $\tau-1$ 的平均有功功率, $\bar{\mathbf{q}}(\tau-1)$ 为在时段 $\tau-1$ 的平均无功功率。时段 τ 的状态 s_τ 为:

$$s_\tau = [(\bar{\mathbf{p}}(\tau-1))^T + j(\bar{\mathbf{q}}(\tau-1))^T, (\mathbf{p}^{\text{ESS}}(\tau-1))^T, (\mathbf{E}^{\text{ESS}}(\tau-1))^T, (\mathbf{n}^{\text{OLTC}}(\tau-1))^T, (\mathbf{y}^{\text{SCR}}(\tau-1))^T]^T \quad (3)$$

3)策略 π :根据时段 τ 的状态 s_τ 采取动作 a_τ 的概率如式(4)所示。

$$\pi(a_\tau | s_\tau) \in [0, 1] \quad (4)$$

4)代价 $c(\tau, t)$:对于平衡配电网,求解时段 τ 内时隙 t 的短时间尺度模型后,若无解或解不满足SOCP精确凸松弛条件,则将节点电压平方的上、下限分别修改为 $+\infty$ 与 0 ,重新求解模型,代价为电压越限惩罚,计算方法如式(5)所示。

$$c(\tau, t) = \frac{100}{N_T} \sum_k (\max(U_k - 1.05, 0) + \max(0.95 - U_k, 0)) \quad (5)$$

式中: U_k 为节点 k 的电压标么值。

若式(5)有解且解满足SOCP精确凸松弛条件,则代价为网损大小,计算方法为:

$$c(\tau, t) = \frac{P_{\text{loss}}(\tau, t)}{N_T} \quad (6)$$

式中: $P_{\text{loss}}(\tau, t)$ 为时段 τ 内时隙 t 的网损。

对于不平衡配电网,求解时段 τ 内时隙 t 的短时间尺度模型后,若无解,则将节点电压平方的上、下限分别修改为 $+\infty$ 与 0 ,重新求解模型,代价为电压越限惩罚,计算方法为:

$$c(\tau, t) = \frac{100}{N_T} \sum_\alpha \sum_k (\max(U_k^\alpha - 1.05, 0) + \max(0.95 - U_k^\alpha, 0)) \quad (7)$$

式中: α 为相位。

若有解,则代价为网损,如式(7)所示。

5)奖励 r_τ :令参数 $c_\tau = \sum_{t=1}^{N_T} c(\tau, t)$,若 $c_\tau > 0.1 \text{ p.u.}$,则

$$r_\tau(s_\tau, a_\tau) = -c_\tau \quad (8)$$

否则

$$r_\tau(s_\tau, a_\tau) = 10 \left(\frac{c_0}{c_\tau} \right)^2 \quad (9)$$

式中: c_0 为常数。

6)长期回报集合 J :MDP的目标是采用最优策略使得长期回报最大。

2.2 基于松弛-预报-校正的DDPG算法

DDPG算法适用于连续动作的场景,配电网中OLTC变比和SCR挡位只能取离散值。因此,不能直接应用DDPG算法,DQN算法不适用于连续动作的场景。ESS充放电功率一般为连续值,DQN算法难以适用。针对此问题,本文将文献[20]中的方法推广至一般情形,即连续-离散动作空间。首先,将离散动作分量松弛为连续动作分量(称为松弛过

程);然后,针对执行器输出的原型动作中对应于 OLTC 变比和 SCR 挡位的分量,在(嵌入)离散动作空间中搜索出 K_m 个最临近点(称为预报过程);最后,每个最邻近点与执行器输出的原型动作中对应于连续动作分量(ESS 充放电功率)组成一个完整动作,并依次输入评价器得到动作价值,选取动作价值最大的动作与环境交互(称为校正过程)。在已知数据集中找出给定点的 K_m 个最临近点能够在对数时间复杂度内完成^[21]。目前,已有大量文献报导该算法。本文采用的基于松弛-预报-校正的改进 DDPG 算法不会大幅增加计算时间,具有很好的可扩展性。

基于松弛-预报-校正的改进 DDPG 算法示意图如附录 C 图 C1(b)所示。本文提出的主动配电网双时间尺度有功无功协调优化程序流程见附录 D 图 D1。

虽然基于 gumble-softmax 重参数化技巧的 DDPG 算法也可以处理离散动作,但是本文方法动作维度等于 OLTC、SCR、ESS 的数量之和,基于 gumble-softmax 重参数化技巧的 DDPG 算法动作维度等于离散动作数量与连续动作数量之和。例如,假设三相平衡配电网含 5 台 OLTC、4 台 SCR、2 台 ESS,每台 OLTC 有 10 个挡位,每台 SCR 有 4 个挡位,则本文方法动作维度为 $5+4+2=11$ 。基于 gumble-softmax 重参数化技巧的 DDPG 算法动作维度为 $10^5 \times 4^4 + 2 = 25\,600\,002$ 。因此,本文方法动作维度远远低于基于 gumble-softmax 重参数化技巧的 DDPG 算法。虽然后者也可以处理离散动作,但其只适用于离散动作数量较少的情形。当离散动作数量很多时,会产生维数灾问题。本文方法的优势之一是不会产生维数灾问题。

为了限制短路电流及便于继电保护的整定与配合,配电网一般采用辐射状运行结构。本文设置 OLTC 与 SCR 动作时间间隔为 1 h,未计及 OLTC、SCR 的动作次数与辐射状网络结构约束,文献[10, 12, 15, 16, 22-25]亦未计及。

本文为实现 SVC 与 DG 逆变器的(毫秒级)快速决策,将 ESS 充放电功率放在长时间尺度进行优化,在一定程度上牺牲了 ESS 的灵活调节能力,但降低了循环次数,提高了寿命。文献[1]也是将 ESS 充放电功率调整时间间隔设置为 1 h,而 SVC 和 DG 无功功率设置为实时调节。

3 仿真算例

3.1 仿真算例 1

IEEE 33 节点配电系统仿真条件如附录 E 所示。算例中,共有 $17 \times 11 \times 11 = 2\,057$ 个离散动作分量。为了验证本文方法的有效性,首先设置

OLTC 变比、SCR 挡位、ESS 充放电功率在每个时段 τ 分别为随机值与固定值。当设为固定值时,OLTC 变比为 1,2 台 SCR 无功补偿均为 0,ESS 充放电功率为 0,只优化求解每个时隙 t 的 SOCP 模型。平均每小时代价的优化结果如附录 F 图 F1 所示。可以看出,每天平均每小时的代价均远大于 0.1 p.u.,意味着每天电压越限均十分严重。第 1 阶段变量取随机值比取固定值导致的电压越限问题更加严重。2 种方法计算时间分别为 1 511 s 和 1 530 s。

本文根据仿真结果代价是否大于 0.1 p.u. 作为判断电压是否越限的标准,原因如下:由于潮流计算的三相功率基准值选取为配电网额定容量(对于 IEEE 33 节点、IEEE 123 节点配电系统分别为 $10\text{ MV}\cdot\text{A}$ 、 $5\text{ MV}\cdot\text{A}$),从式(5)至式(9)可知,若电压未越限,则代价为每个时段的平均网损(有功损耗),其值必远小于 0.1 p.u.。若大于 0.1 p.u.,则表明网损率远大于 10%,这在实际配电网中是不可能的。每天平均每小时的代价均远大于 0.1 p.u.,意味着每天的电压越限均十分严重,这是由于代价计及了电压越限惩罚。

将 OLTC 变比设置为连续变量,并设定 OLTC 变比、SCR 挡位、ESS 充放电功率在每个时隙 t 均可调节,构建 288 个时隙单一短时间尺度传统日前混合整数 SOCP 模型。基于 MATLAB 平台采用 Mosek 9.1.4 软件包求解,为提高计算速度,设置相对对偶间隙为 $+\infty$,平均每小时最优代价为 0.009 p.u.,如图 1 中绿色点线所示。

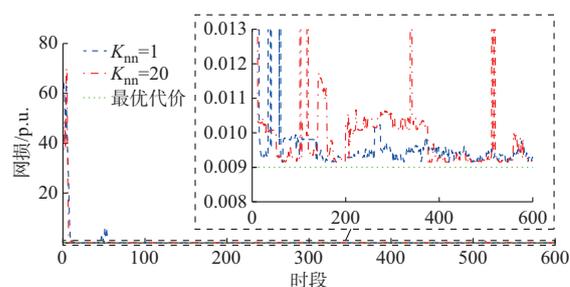


图 1 算例 1 中 IEEE 33 节点配电网平均每小时的代价曲线
Fig. 1 Curves of average cost per hour of IEEE 33-bus distribution network in case 1

采用本文方法,针对执行器输出的动作,在离散动作分量空间搜索出 $K_m = 1$ 与 $K_m = 20$ 时最邻近的动作分量,优化结果分别如图 1 中蓝色虚线与红色点划线所示。可以看出,采用本文方法,起始阶段平均每小时代价很高,这是因为 DDPG 算法在起始阶段执行器采用的是随机动作策略。在第 60 d 后(训练了 $60 \times 24 \text{步} = 1\,440 \text{步}$),网损开始变得很

低,接近最优值0.009 p.u.。

$K_{mn}=1$ 与 $K_{mn}=20$ 时的总优化计算时间分别为1 307 s和1 487 s,对应每日数据的平均优化计算时间分别为2.178 s和2.478 s。求解传统288个时隙日前规划混合整数SOCP计算时间为114 s。当 $K_{mn}=20$ 时,本文方法的计算速度约为传统方法的46倍,每个时隙 t 的平均计算时间为0.008 6 s。因此,即使每个时隙 t 持续时间设置为1 s,本文所提方法仍满足实时控制需求。

3.2 仿真算例2

设置每个DG最大功率和容量分别为1 MW和1 MV·A, c_0 为0.005 p.u.。其他仿真条件与算例1相同。训练过程中的平均每小时代价优化结果如图2所示。可以看出,采用本文方法,当 $K_{mn}=20$ 时,在第70 d后(训练了 70×24 步=1 680步),平均每小时代价十分接近传统288个时隙单一短时间尺度混合整数SOCP模型优化结果的最优网损0.003 8 p.u.。 $K_{mn}=20$ 时,前100个时段电压频繁越限的原因是DDPG算法在起始阶段执行器采用的是随机动作策略。当任务较困难时,改进DDPG智能体需要训练足够长的步数才能学习到最优策略。然而,当 $K_{mn}=1$ 时,训练过程很不平稳,电压越限问题频繁发生。这是因为风电比例很高(风电比例是算例1的10倍),部分支路存在反向潮流,电压分布范围较宽。为避免电压与电流越限,OLTC变比、SCR挡位与ESS充放电功率必须精准设定。然而,在离散动作空间取1个最临近点不能保证解的可行性。因此,任务很困难时,选取 $K_{mn}=1$ 是不合适的。

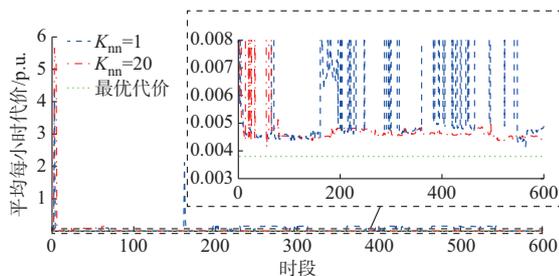


图2 算例2中IEEE 33节点配电网平均每小时的代价曲线
Fig. 2 Curves of average cost per hour of IEEE 33-bus distribution network in case 2

3.3 仿真算例3

设置OLTC最大、最小变比分别为1.1与0.9。其他仿真条件与算例1相同。此算例中共有 $33 \times 11 \times 11 = 3993$ 个离散动作分量。若OLTC变比不在区间 $[0.95, 1.05]$ 内,则OLTC二次侧电压越限。相比于算例1,此算例中DDPG智能体找到(近

似)最优解要困难得多,平均每小时代价优化结果如图3所示。可以看出,采用本文方法,当 $K_{mn}=40$ 时,在第250 d后(训练了 250×24 步=6 000步),平均每小时代价十分接近传统288个时隙单一短时间尺度混合整数SOCP模型优化结果的最优网损0.008 8 p.u.。然而,当 $K_{mn}=400$ 时,训练过程中的平均每小时代价一直很高,算法不收敛。这是因为最临近点选取得太多导致DDPG智能体没有得到有效训练,神经网络参数未进化,一直找不到可行解。因此,当任务很困难时, K_{mn} 选取得过大也是不合适的。此外,当 $K_{mn}=1$ 时,训练过程很不平稳,电压越限问题频繁发生。这是因为在离散动作空间取1个最临近点不能保证解的可行性。这再次证明,任务很困难时,选取 $K_{mn}=1$ 是不合适的。

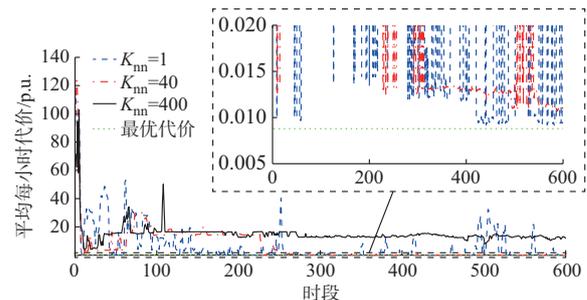


图3 算例3中IEEE 33节点配电网平均每小时的代价曲线
Fig. 3 Curves of average cost per hour of IEEE 33-bus distribution network in case 3

3.4 仿真算例4

IEEE 123节点配电系统仿真条件如附录E所示。假设根节点电压固定为1.05 p.u.。此时,离散动作分量数量为 $5^3 \times 2^6$ 个=8 000个。设置OLTC变比、电容器是否投入、ESS充放电功率在每个时段 τ 分别为随机值与固定值。当设为固定值时,电容器均不投入,ESS充放电功率为0,位于支路9-14、25-26、119-67的OLTC变比分别为1.000、1.000、1.025,只优化求解每个时隙 t 的二次规划模型,平均每小时代价优化结果如附录F图F2所示。可以看出,每天平均每小时的代价均远大于0.1 p.u.,意味着电压越限均十分严重。而且,第1阶段变量取固定值比取随机值导致的电压越限问题更加严重。

采用本文所提方法,当 K_{mn} 为1与20时,平均每小时代价优化结果分别如图4中蓝色虚线与红色点划线所示。在第20 d后(训练了 20×24 步=480步),当 $K_{mn}=20$ 时,平均每小时代价接近最优值0.018 8 p.u.。

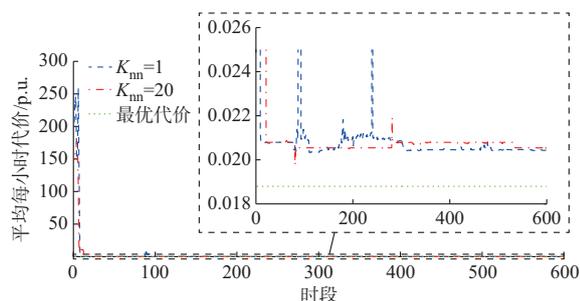


图4 算例4中IEEE 123节点配电网平均每小时的代价曲线

Fig. 4 Curves of average cost per hour of IEEE 123-bus distribution network in case 4

3.5 仿真算例5

假设根节点接入一台三相OLTC,最大、最小变比分别为1.05与0.95,步长为0.025。其他仿真条件与算例4相同。此算例中共有 $5^4 \times 2^6$ 个=40 000个离散动作分量。当第10阶段变量取随机值或固定值当设定为固定值时,电容器均不投入,ESS充放电功率为0,位于支路123-1、9-14、25-26、119-67的OLTC变比分别为1.000、1.050、0.950、1.050。平均每小时代价优化结果如附录F图F3所示。可以看出,相比于附录F图F2,电压越限问题更加严重。而且,第10阶段变量取随机值比取固定值导致的电压越限问题更加恶化。这是因为根节点OLTC变比对电压越限的影响最大。2种方法优化计算总时间分别为7 231 s与7 270 s。

采用本文所提方法,平均每小时代价优化结果如图5所示。可以看出,在第238 d后(训练了 238×24 步=5 712步),当 $K_m=20$ 时,平均每小时代价接近最优值0.017 9 p.u.,训练过程平稳度与收敛速度远高于 $K_m=1$ 时。

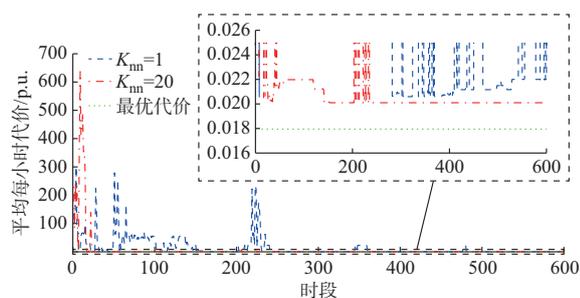


图5 算例5中IEEE 123节点配电网平均每小时的代价曲线

Fig. 5 Curves of average cost per hour of IEEE 123-bus distribution network in case 5

当 K_m 为1与20时,总优化计算时间分别为629 4 s和6 547 s,每日数据的平均优化计算时间分别为10.490 0 s和10.911 7 s。将相对对偶间隙设置为无穷大,求解传统144个时隙单一短时间尺度日

前规划混合整数SOCP计算时间为199 s。当 $K_m=20$ 时,本文方法计算速度约为传统方法的18.237倍,每个时隙 t 的平均计算时间为0.075 8 s。因此,即使时隙 t 设置为1 s,本文所提方法仍满足实时控制的需求。

3.6 与现有多智能体DQN算法比较

对比图5与文献[16]可以看出,采用本文方法训练过程收敛速度与平稳度远高于文献[16]中的多智能体DQN算法。这是因为多智能体协调探索与利用比单智能体复杂与困难得多。文献[16]将连续变量离散化,离散动作数量呈指数增长,而且构建的单一长时间尺度MDP未计及不同类型可调设备动作速度差异性,降低了调度灵活性与最优性。

4 结语

本文方法的突出优点是很容易在实际配电网中实现。基于IEEE 33节点与IEEE 123节点配电系统仿真结果表明,针对执行器输出的原型动作,在离散动作分量空间选取的最临近点数量 K_m 对训练过程收敛速度与平稳度具有较大影响。当任务较困难时, K_m 太大或太小,如 K_m 为1或400,可能导致训练过程很不平稳或不收敛。 K_m 的取值适中,如20或40,即可使得训练过程较平稳。本文方法优化结果十分接近于OLTC、SCR与ESS参与短时间尺度调节的多时隙单一短时间尺度日前混合整数SOCP或二次规划优化结果。然而,当 $K_m=20$ 时,本文方法计算速度是其18~42倍。而且,本文方法DRL训练过程收敛速度与平稳度远高于现有单一慢时间尺度多智能体DQN算法。

本文未计及电动汽车的充放电功率约束与调节作用,未采用测试集验证基于松弛-预报-校正的DDPG算法泛化能力。进一步研究的工作重点是在短时间尺度优化模型中计及电动汽车充放电功率约束与调节作用,以及采用测试集验证基于松弛-预报-校正的DDPG算法泛化能力。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

参考文献

- [1] ZHANG J, CUI M J, HE Y G, et al. Multi-period two-stage robust optimization of radial distribution system with cables considering time-of-use price [J]. Journal of Modern Power Systems and Clean Energy, 2023, 11(1): 312-323.
- [2] 王泉,何恰刚,马恒瑞,等.考虑规模化储能的配电网电压分布式控制[J].电力自动化设备,2022,42(2):25-30.

- WANG Xiao, HE Yigang, MA Hengrui, et al. Distributed voltage control of distribution network considering large-scale energy storage [J]. *Electric Power Automation Equipment*, 2022, 42(2): 25-30.
- [3] 刘其辉, 逢思敏, 吴林林, 等. 大规模风电汇集系统电压不平衡机理、因素及影响规律[J]. *电工技术学报*, 2022, 37(21): 5435-5450.
- LIU Qihui, PANG Simin, WU Linlin, et al. The mechanism, factors and influence rules of voltage imbalance in wind power integration areas [J]. *Transactions of China Electrotechnical Society*, 2022, 37(21): 5435-5450.
- [4] 徐博涵, 向月, 潘力, 等. 基于深度强化学习的含高比例可再生能源配电网就地分散式电压管控方法[J]. *电力系统保护与控制*, 2022, 50(22): 100-109.
- XU Bohan, XIANG Yue, PAN Li, et al. Local decentralized voltage management of a distribution network with a high proportion of renewable energy based on deep reinforcement learning [J]. *Power System Protection and Control*, 2022, 50(22): 100-109.
- [5] 冯昌森, 张瑜, 谢路耀, 等. 配电系统双时间尺度电压管理的深度强化学习方法[J]. *电力系统自动化*, 2022, 46(12): 202-209.
- FENG Changsen, ZHANG Yu, XIE Luyao, et al. Deep reinforcement learning approach for dual-timescale voltage management in distribution system [J]. *Automation of Electric Power Systems*, 2022, 46(12): 202-209.
- [6] 习伟, 李鹏, 李鹏, 等. 基于深度强化学习的分布式电源就地自适应电压控制方法[J]. *电力系统自动化*, 2022, 46(22): 25-31.
- XI Wei, LI Peng, LI Peng, et al. Adaptive local voltage control method for distributed generator based on deep reinforcement learning [J]. *Automation of Electric Power Systems*, 2022, 46(22): 25-31.
- [7] 蒲天骄, 杜帅, 李焯, 等. 面向隐私保护基于联邦强化学习的分布式电源协同优化策略[J]. *电力系统自动化*, 2023, 47(8): 62-70.
- PU Tianjiao, DU Shuai, LI Ye, et al. Collaborative optimization strategy of distributed generators based on federated reinforcement learning for privacy preservation [J]. *Automation of Electric Power Systems*, 2023, 47(8): 62-70.
- [8] 胡丹尔, 彭勇刚, 韦巍, 等. 多时间尺度的配电网深度强化学习无功优化策略[J]. *中国电机工程学报*, 2022, 42(14): 5034-5044.
- HU Daner, PENG Yonggang, WEI Wei, et al. Multi-time scale reactive power optimization strategy for deep reinforcement learning in distribution network [J]. *Proceedings of the CSEE*, 2022, 42(14): 5034-5044.
- [9] WANG W, YU N P, SHI J, et al. Volt-var control in power distribution systems with deep reinforcement learning [C]// 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), October 21-23, 2019, Beijing, China: 1-7.
- [10] XU H C, DOMÍNGUEZ-GARCÍA A D, SAUER P W. Optimal tap setting of voltage regulation transformers using batch reinforcement learning [J]. *IEEE Transactions on Power Systems*, 2020, 35(3): 1990-2001.
- [11] AL-SAFFAR M, MUSILEK P. Reinforcement learning-based distributed BESS management for mitigating overvoltage issues in systems with high PV penetration [J]. *IEEE Transactions on Smart Grid*, 2020, 11(4): 2980-2994.
- [12] 李琦, 乔颖, 张宇精. 配电网持续无功优化的深度强化学习方法[J]. *电网技术*, 2020, 44(4): 1473-1480.
- LI Qi, QIAO Ying, ZHANG Yujing. Deep reinforcement learning method for continuous reactive power optimization of distribution network [J]. *Power System Technology*, 2020, 44(4): 1473-1480.
- [13] WANG W, YU N P, GAO Y Q, et al. Safe off-policy deep reinforcement learning algorithm for volt-var control in power distribution systems [J]. *IEEE Transactions on Smart Grid*, 2020, 11(4): 3008-3018.
- [14] CAO D, HU W H, ZHAO J B, et al. A multi-agent deep reinforcement learning based voltage regulation using coordinated PV inverters [J]. *IEEE Transactions on Power Systems*, 2020, 35(5): 4120-4123.
- [15] 倪爽, 崔承刚, 杨宁, 等. 基于深度强化学习的配电网多时间尺度在线无功优化[J]. *电力系统自动化*, 2021, 45(10): 77-85.
- NI Shuang, CUI Chenggang, YANG Ning, et al. Multi-timescale online optimization for reactive power of distribution network based on deep reinforcement learning [J]. *Automation of Electric Power Systems*, 2021, 45(10): 77-85.
- [16] ZHANG Y, WANG X N, WANG J H, et al. Deep reinforcement learning based volt-var optimization in smart distribution systems [J]. *IEEE Transactions on Smart Grid*, 2021, 12(1): 361-371.
- [17] 刘朝阳, 穆朝絮, 孙长银. 深度强化学习算法与应用研究现状综述[J]. *智能科学与技术学报*, 2020, 2(4): 314-326.
- LIU Zhaoyang, MU Chaoxu, SUN Changyin. An overview on algorithms and applications of deep reinforcement learning [J]. *Chinese Journal of Intelligent Science and Technology*, 2020, 2(4): 314-326.
- [18] 孙长银, 穆朝絮. 多智能体深度强化学习的若干关键科学问题 [J]. *自动化学报*, 2020, 46(7): 1301-1312.
- SUN Changyin, MU Chaoxu. Important scientific problems of multi-agent deep reinforcement learning [J]. *Acta Automatica Sinica*, 2020, 46(7): 1301-1312.
- [19] ZHANG L, WANG G, GIANNAKIS G B. Real-time power system state estimation and forecasting via deep unrolled neural networks [J]. *IEEE Transactions on Signal Processing*, 2019, 67(15): 4069-4077.
- [20] DULAC-ARNOLD G, EVANS R, HASSELT H V, et al. Deep reinforcement learning in large discrete action spaces [EB/OL]. [2022-12-15]. <https://arxiv.org/abs/1512.07679>.
- [21] MUJA M, LOWE D G. Scalable nearest neighbor algorithms for high dimensional data [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(11): 2227-2240.
- [22] ZHANG J, CUI M J, LI B, et al. Fast solving method based on linearized equations of branch power flow for coordinated charging of EVs (EVCC) [J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(5): 4404-4418.
- [23] ZHANG J, CUI M J, HE Y G. Multi-period fast robust optimization for partial distributed generators (DGs) providing ancillary services [J]. *Energies*, 2021, 14(16): 4911.

- [24] ZHANG J, HE Y G. Fast solving method for two-stage multi-period robust optimization of active and reactive power coordination in active distribution networks[J]. IEEE Access, 2022, 11: 30208-30222.
- [25] DING T, LIU S Y, YUAN W, et al. A two-stage robust reactive power optimization considering uncertain wind power integration in active distribution networks [J]. IEEE Transactions on Sustainable Energy, 2016, 7(1): 301-311.

向:电力系统建模、主动配电网技术、电动汽车有序充电等。
E-mail:z_jj1219@sina.com

崔明建(1987—),男,博士,教授,主要研究方向:风力预测、机组组合、配电网物理信息系统等。E-mail:mingjian.cui@ieee.org

姚潇毅(1990—),男,硕士,工程师,主要研究方向:电动汽车有序充电、电力规划与设计等。E-mail:yaoxy041x@163.com

张 剑(1982—),男,通信作者,博士,讲师,主要研究方

(编辑 杨松迎)

Dual-timescale Active and Reactive Power Coordinated Optimization for Active Distribution Network Based on Data-driven and Physical Model

ZHANG Jian¹, CUI Mingjian², YAO Xiaoyi³, HE Yigang⁴

(1. School of Electrical and Automation Engineering, Hefei University of Technology, Hefei 230009, China;

2. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China;

3. Bengbu Electric Power Supply Company of State Grid Anhui Electric Power Company, Bengbu 233000, China;

4. School of Electrical and Automation, Wuhan University, Wuhan 430072, China)

Abstract: When a high proportion of intermittent distributed generators (DGs) and electric vehicles are connected to the distribution network, it is easy to cause frequent, fast and dramatic fluctuations in power and voltage. This paper combines both the data-driven and physical modeling approaches to propose a coordinated optimal strategy for dual-timescale active and reactive power in distribution networks. For the short-timescale (minute or second) power fluctuations, a second-order cone programming (SOCP) model and a quadratic programming model are constructed for balanced and unbalanced distribution networks, respectively, with static var compensator (SVC) and reactive power of DG as decision variables and network loss minimization as the objective function, taking into account physical constraints. For the long-timescale (hour scale) optimization, the Markov decision process (MDP) is constructed with the tap ratio of on-load transformer changers (OLTCs), the tap position of switchable capacitors reactors (SCRs), and the charging/discharging power of energy storage systems (ESSs) as the actions, network loss as the cost, taking into account the crossing penalty of bus voltage. To overcome dimension curses in continuous-discrete action space, this paper uses a relaxation-prediction-correction based deep deterministic policy gradient (DDPG) reinforcement learning algorithm. The effectiveness of the proposed method is verified by IEEE 33-bus and IEEE 123-bus distribution systems.

This work is supported by National Natural Science Foundation of China (No. 52207130).

Key words: active distribution network; distributed generator (DG); deep reinforcement learning; second-order cone programming (SOCP); quadratic programming (QP)

