

基于强化学习算法的微电网优化策略

李子凯¹, 杨波¹, 周忠堂¹, 张健², 徐明珠¹

(1. 国网山东省电力公司临沂供电公司, 山东 临沂 2760001; 2. 国网(山东)电动汽车服务有限公司, 山东 济南 250000)

摘要:分布式能源具有小规模波动和间歇性的特点,导致微电网运行策略难以制定。微电网有效集成多种分布式能源和外部电网,多能源微电网管理正成为一项非常复杂的任务。针对该问题,提出一种在负荷需求、可再生能源和储能设备等综合因素影响下的微电网实时优化运行策略。该策略首先基于强化学习框架,将微电网运行问题建模为马尔可夫决策过程,然后以最小化微电网电压波动和运行损耗为目的构建微电网策略优化模型。为有效利用微电网的互联结构,在近端策略优化算法的基础上,设计一种图注意力近端策略优化算法(graph attention proximal policy optimization, GT-PPO),该算法使用注意力机制和图神经网络学习微电网节点的相关性,以学习各类环境下不同时段多能源微电网最优调度策略。最后,采用改进的IEEE 33节点、IEEE 118节点两种规格的微电网进行仿真实验。实验结果表明,该优化策略可以实现微电网的实时优化,且结果优于传统的近端策略优化(proximal policy optimization, PPO)算法和双深度Q网络(double deep Q network, DDQN)算法。

关键词:微电网;近端策略优化;图自注意力网络;策略优化

中图分类号:TP18

文献标识码:A

文章编号:1007-9904(2024)06-0027-09

Optimization Strategy for Microgrid Based on Reinforcement Learning Algorithm

LI Zikai¹, YANG Bo¹, ZHOU Zhongtang¹, ZHANG Jian², XU Mingzhu¹

(1.State Grid Linyi Power Supply Company, Linyi 276001, China;

2.State Grid Shandong EV Service Co., Ltd., Jinan 250000, China)

Abstract: Distributed energy has the characteristics of small-scale fluctuations and intermittency, making it difficult to formulate operational strategies for microgrids. As an effective way to integrate multiple distributed energy sources and external grids, multi-energy microgrid management is becoming a very complex task. A microgrid real-time optimal operation strategy was proposed under the influence of comprehensive factors such as load demand, renewable energy sources and energy storage devices. Firstly, based on the reinforcement learning framework, the microgrid operation problem was modeled as a Markov decision process, and then a microgrid strategy optimization model was constructed with the aim of minimizing voltage fluctuations and operational losses in the microgrid. In order to effectively utilize the interconnection structure of the distribution network, a graph attention proximal policy optimization (GT-PPO) algorithm was designed on the basis of the proximal policy optimization algorithm. This algorithm uses an attention mechanism and a graph neural network to learn the correlation of distribution network nodes to formulate the optimal scheduling strategy for multi-energy distribution networks at different times under various environments. Simulation experiments were conducted using two specifications of the improved IEEE 33 node and IEEE 118 node distribution networks. The experimental results show that the optimization strategy can achieve real-time optimization of microgrids, and the results are better than the traditional proximal policy optimization (PPO) algorithm and double deep Q network (DDQN) algorithm.

Keywords: microgrid; proximal policy optimization; graph attention networks; strategy optimization

基金项目:国网山东省电力公司科技项目资助项目“电动两轮车换电业务建设运营关键技术研究”(520607220008)。

State Grid Shandong Electric Power Company Science and Technology Project Funding Project "Research on Key Technologies for the Construction and Operation of Electric Two wheeled Vehicle Power Exchange Business Fund"(520607220008).

0 引言

微电网是一种新型的电力系统,由多个分布式电源、负荷和存储装置组成的小型电力系统^[1]。为保证电力输送的安全可靠,需要电力控制室的人工专家持续进行检测和管理。但随着分布式光伏发展,装机容量逐年提升,电动车负荷大量增多,这些因素给低压微电网带来很大影响,多能源微电网管理正成为一项非常复杂的任务^[2]。微电网的能源管理主要面临优化调度问题。在分布式能源干预的情况下,分布式能源的小规模波动性和间歇性特点的随机性给微电网带来了许多优化分配和调度问题^[3]。分布式能源的随机性使得微电网功率平衡变得困难,从而引起微电网电压波动。

微电网中的电压上升和波动问题已经被研究了二十多年^[4]。传统的电压调节装置,如有载分接开关和电容器组^[5],通常安装在变电站处,因此可能无法有效调节线路远端的电压。最先进的有功电压控制策略大致可分为两类:基于最优潮流的无功功率调度^[6-7]、基于本地电压和功率测量的下垂控制^[8-9]。最优潮流方法的主要限制是需要精确的系统模型^[10],此外,解决约束优化问题很耗时,因此很难对负载分布的快速变化做出反应^[11]。下垂控制仅取决于其局部测量,但其性能依赖于手动设计的参数,并且由于缺乏全局信息,通常是次优的。

在微电网多能源协调控制方面。刘一欣等人提出一种两阶段鲁棒优化方法实现了系统日运行成本最小化^[12]。戴涛等人使用模糊控制方法对微电网进行管理,提高微电网可靠性^[13]。上述研究工作都是依赖传统优化方法,传统的能量优化算法都是基于数值优化的算法,算法复杂度高且响应速度慢。

随着强化学习和量测技术研究的进一步深入,已有许多研究在微电网管理和控制中使用强化学习方法,并取得较好的效果。张沛等考虑新能源和负荷的时变特性,将无功优化问题构建成强化学习问题,使用近端策略优化算法求解^[14]。赵鹏杰等将微电网优化运行问题转化为马尔可夫决策过程,使用深度确定性策略梯度算法求解经济利益和居民满意度最大时微电网运行策略^[15]。Deunsol等人提出一种基于非策略的actor-critic算法的电网优化运行模型,采用分层策略和后状态表示^[16]。刘俊峰等提出

一种微能源电网能量管理模型,使用深度Q网络(deep Q network, DQN)算法求解能量管理策略^[17]。

针对微电网中源网荷储多因素的互动问题,考虑微电网运行的现状,研究储能参与微电网运行的能量互动机制以及与微电网内光伏、风机、用户负荷等进行协同调控的模式与策略。为有效利用微电网的互联结构,在近端策略优化算法的基础上,设计一种图自注意力近端策略优化(graph attention proximal policy optimization, GT-PPO)算法,以学习各类环境下不同时段多能源微电网最优调度策略。

GT-PPO算法通过图神经网络对电网的互连结构加以利用,进一步,利用邻接矩阵 M 屏蔽节点的注意力权重,采用图注意力网络(graph attention network, GAT)计算节点注意力系数,使得近端策略优化(proximal policy optimization, PPO)算法能够关注更重要的节点信息。

1 微电网系统及优化模型

1.1 微电网结构

微电网模型结构如图1所示,微电网模型包括风力发电、光伏发电、储能电池、负荷、外部电网接口以及微电网母线部分。建立的模型通过控制各节点输入输出功率大小,实现微电网稳定经济运行。

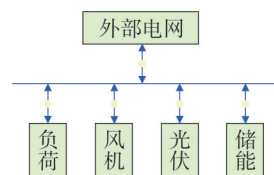


图1 多能源微电网模型

Fig.1 Multi-energy distribution network model

电网的状态由各种特征组成,例如拓扑配置(每条电力线的连通性和每个变电站中的总线分配),分布式电源提供的电量、每个负载所需的电量、每条线路传输的电量等等。分布式电源提供的功率和负载的功率随时间变化。

1.2 微电网主要设备的数学模型

1.2.1 储能设备

储能设备在微电网能量过剩时储存能量,并在能量不足时提供能量。在储能设备模型中,储能设

备作为短期储备,其充放电响应信号由电池发送。微电网中的储能设备可以维持微电网能源的供需平衡。电池在 t 时刻的实时能量为

$$S_{\text{bat}}(t) = \begin{cases} S_{\text{bat}}(t-1) + \eta_b \int P_{\text{bat}}^{\text{cha}} dt \\ S_{\text{bat}}(t-1) - \int \frac{P_{\text{bat}}^{\text{dis}}}{\eta_b} dt \end{cases} \quad (1)$$

式中: $S_{\text{bat}}(t-1)$ 为充电和放电之前的能量; $P_{\text{bat}}^{\text{cha}}$ 和 $P_{\text{bat}}^{\text{dis}}$ 分别为充电功率和放电功率; η_b 为充电和放电效率。

单个储能设备充放电响应行为受电池容量约束。电池具有一定的续航里程限制,如式(2)一式(4)所示。

$$S_{\text{bat},\min} \leq S_{\text{bat}}(t) \leq S_{\text{bat},\max} \quad (2)$$

$$0 \leq P_{\text{dis},t} \leq P_{\text{dis},\max} \quad (3)$$

$$0 \leq P_{\text{cha},t} \leq P_{\text{cha},\max} \quad (4)$$

式中: $S_{\text{bat},\min}$ 和 $S_{\text{bat},\max}$ 分别为电池存储容量的下限和上限; $P_{\text{dis},\max}$ 和 $P_{\text{cha},\max}$ 分别为电池放电和充电的最大功率; $P_{\text{dis},t}$ 和 $P_{\text{cha},t}$ 分别为时刻电池放电和充电的实时功率。

1.2.2 电网母线

为实现可再生能源的完全消纳,默认风机、光伏等微电网内分布式电源功率全部并网。微电网在能量传输时有损耗,基于微电网需求侧的稳定运行,母线功率必须满足负载所需^[18],可建模为:

$$P_t^{\text{grid}} + P_t^{\text{PV}} + P_t^{\text{WT}} + P_t^{\text{ES}} \geq P_t^{\text{L}} \quad (5)$$

$$P_t^{\text{ES}} = P_t^{\text{dis}} - P_t^{\text{ch}} \quad (6)$$

式中: P_t^{grid} 为 t 时刻外部电网输入多能源微电网功率; P_t^{PV} 为 t 时刻光伏电源输出功率; P_t^{WT} 为 t 时刻风机输出功率; P_t^{ES} 为 t 时刻储能输出功率,若储能处于充电状态,则为负值; P_t^{dis} 和 P_t^{ch} 分别为 t 时刻储能放电和充电功率。

1.2.3 可再生能源发电和负荷

可再生能源发电和负荷在实际微电网中随机波动,在 t 时刻微电网的净负载 P_t^{net} 为

$$P_t^{\text{net}} = P_t^{\text{L}} - P_t^{\text{PV}} - P_t^{\text{WT}} \quad (7)$$

考虑到随机性,微电网的净负载的序列被公式化为离散时间随机过程。

1.2.4 微电网优化模型

微电网运营商通过求解以下优化问题来确定运行方案,使得电压波动和负载运行损耗最小,如式

(8)所示。

$$\min \left[\sum_{t=1}^T \left(\frac{V_t - V}{V} + \frac{P_t^{\text{grid}} - P_t^{\text{net}}}{P_t^{\text{net}}} \right) \right] \quad (8)$$

式中: V 为微电网额定电压; V_t 为 t 时刻微电网节点实际电压; P_t^{grid} 和 P_t^{net} 分别为 t 时刻外部电网输入功率和微电网净负载。

以式(8)为目标函数,式(1)一式(7)为约束条件,求解微电网在电压波动最小且运行损耗最低时的运行策略,形成混合整数二次规划(mixed integer quadratic programming, MIQP)问题。

1.3 马尔可夫决策过程

在强化学习中,智能体是一个基于学习算法的决策系统,用于在环境中进行交互和学习。使智能体在特定环境下不断进行动作试错,根据不同环境状态输出不同时刻动作,环境根据动作确定下一时刻状态,智能体根据状态动作的奖励求解不同环境下的最优响应动作^[19]。马尔可夫决策过程指在强化学习任务中的环境信息具有马尔可夫属性^[20]。使用五元组 (S, A, P, r, γ) 定义的马尔可夫决策过程表示强化学习任务,其中: S 为环境状态空间,设 s_t 为 t 时刻的环境状态; A 为动作空间,设 a_t 为 t 时刻智能体的动作; P_t 为状态转移概率; r 为即时奖励; γ 为奖励折扣因子。

若折扣因子为 γ ,前 T 时段,累积奖励定义为

$$R_t = \sum_{t=0}^{T-1} \gamma^t r_{t+1} \quad (9)$$

强化学习的任务是找到最佳策略 π ,在该策略下,所有状态的预期回报最大化,为

$$\pi^* = \operatorname{argmax}_{\pi} E_{\pi} [R_t | \pi] \quad (10)$$

式中: R_t 为 t 时刻的累积奖励; argmax 为期望最大化时对应的参数值的函数; E_{π} 为期望值。

马尔可夫决策过程为智能体根据策略在状态 s_t 下选择动作 a_t 并且状态以 $p_{\pi}(a_t | s_t)$ 的概率转移到下一个时刻状态 s_{t+1} ,同时得到环境的反馈奖励 $r(s_t, a_t)$ 。将微电网策略优化问题转化为强化学习任务的关键是进行马尔可夫决策的建模。

1.3.1 状态空间

微电网优化运行策略中,状态空间 S 由强化学习所需电网环境状态信息集合组成。环境中状态信息包括储能状态,风机光伏等分布式电源出力,母线

电压和负载功率。因此,时刻 t 处的状态量定义为

$$s_t = \{P_{i,t}^L, Q_{i,t}^L, P_{i,t}^W, P_{i,t}^{PV}, E_{i,t}, V_{i,t}\} \quad (11)$$

式中: $P_{i,t}^L, Q_{i,t}^L$ 分别为 t 时刻节点 i 的负载有功功率和无功功率的值; $P_{i,t}^W$ 为 t 时刻节点 i 风机的有功功率; $P_{i,t}^{PV}$ 为 t 时刻节点 i 光伏的有功功率; $E_{i,t}$ 为 t 时刻节点 i 储能装置的能量; $V_{i,t}$ 为 t 时刻节点 i 的电压。

1.3.2 动作空间

代理每次获得状态信息时,都会根据策略在动作空间中进行选择。在多能源微电网中,控制动作包括分布式电源功率大小、变电站输出功率大小和储能充放电功率大小。智能体动作空间为所有动作的集合。时刻 t 处的动作值定义为

$$a_t = \{P_{i,t}^{ET}, Q_{i,t}^{ET}, P_{i,t}^W, P_{i,t}^{PV}\} \quad (12)$$

式中: $P_{i,t}^{ET}$ 为 t 时刻节点 i 储能装置的有功功率; $Q_{i,t}^{ET}$ 为 t 时刻节点 i 储能装置的无功功率; $P_{i,t}^W$ 为 t 时刻节点 i 风机的有功功率; $P_{i,t}^{PV}$ 为 t 时刻节点 i 光伏的有功功率。

1.3.3 奖励函数

在强化学习中,奖励值代表策略好坏的量化,环境根据智能体选择动作生成下一时刻的状态,并根据下一时刻的状态给予奖励,智能体会通过逐渐约束动作满足动作变化率的方法获得最大奖励^[21]。微电网的管理目标在该策略下是电网电压最稳定,运行损耗最低。奖励函数定义为

$$r_t = -[P_{\text{loss}}(t) + \varphi] \quad (13)$$

式中: $P_{\text{loss}}(t)$ 为 t 时刻微电网的有功损耗; φ 为电压波动时的惩罚项。

1.3.4 目标函数

多能源微电网优化运行策略的目标是找到最佳的运行策略,以最大化 T 个时间段内累计奖励的期望,目标函数为

$$\max_{\pi \in \Pi} J(\pi) = E_{\pi} \left[\sum_{t=1}^T \gamma^{t-1} r(t) \right] \quad (14)$$

其中: Π 为所有策略的集合; $E_{\pi}(\cdot)$ 为在策略 π 时随机变量的期望值。

电网在给定的时间段内进行模拟,通常以5 min为间隔模拟。当智能体无法管理网格时,模拟可能会过早地终止,即负载所需的电量没有完全供应,或者由于应用某个动作而导致某个负载与电网断开连接。因此,智能体的目标是通过最小化故障惩罚和

功率损失惩罚来安全有效地运行电网。

2 基于PPO的微电网策略优化

介绍强化学习算法中的PPO和图自注意力网络,并提出了基于图自注意力网络的GT-PPO算法应用于微电网策略优化。

2.1 PPO

PPO算法是一种深度强化学习算法,结合了深度学习的计算能力和强化学习的学习能力的优势,具有性能优良和适应性强的特点^[22]。PPO算法继承Actor-Critic算法框架,通过Actor网络进行动作选择,Critic网络进行状态值估计,最后通过策略梯度方法进行网络更新。PPO能够解决动作空间连续的学习问题,且PPO算法作为TRPO算法的进阶版,计算方式更加简便。

由于TRPO的计算过程非常复杂,每一步更新的运算计算较为复杂,同时随着维度的提高,计算量大幅度增加,PPO对此进行了简化^[23]。计算公式为

$$\text{maximize } \hat{\mathbb{E}}_t \left[\min \left(r_t(\theta) \hat{A}_{\pi_{\theta,t}}, \text{clip} \left(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon \right) \hat{A}_{\pi_{\theta,t}} \right) \right] \quad (15)$$

式中: $\text{clip}(x, l, r) = \max(\min(x, r), l)$ 为剪裁函数,表示将 x 变化控制在 $[l, r]$ 内; $r_t(\theta)$ 为新旧策略比率; $\hat{A}_{\pi_{\theta,t}}$ 为优势函数 $A_{\pi_{\theta}}$ 的采样估计。PPO算法继承了TRPO算法的稳定可靠特点,又进一步降低运算难度,算法的整体性能得到了提高^[24]。PPO算法的Actor网络和Critic网络的参数更新如式(16)一式(18)所示。

$$w_{t+1} = w_t + \alpha^w \delta_t \nabla V(s_t, w_t) \quad (16)$$

$$\delta_t = R_t + \gamma V(s_{t+1}, w_{t+1}) - V(s_t, w_t) \quad (17)$$

$$\theta_{t+1} = \theta_t + \alpha^{\theta} \delta_t \nabla \log \pi_{\theta_t}(a_t | s_t) \quad (18)$$

式中: w_t 和 θ_t 为 t 时刻Critic网络和Actor网络的参数; α^w 和 α^{θ} 分别为两个网络参数的更新步长; δ_t 为时序差分误差值,表示参数更新的方向和大小; R_t 为 t 时刻下的环境奖励。

在PPO算法中,轨迹缓存池收集每次智能体与环境的交互轨迹,在完成一次策略更新后释放全部轨迹^[25]。动作网络和评价网络使用缓存池中的轨迹信息进行训练,进行动作网络和评价网络参数更新,更新后的网络具有更准确的价值评估和动作选择

性能^[26]。

2.2 图注意力

目前,神经网络的研究已经趋于成熟。由于不同种类的神经网络对输入数据的特征提取能力也有所不同,在选择神经网络结构时,需要考虑数据特征和任务类型。引入注意力机制,对电网的图结构特征加以利用。通过电力设备的连接情况,计算每个电力设备节点与相邻设备节点的相关性,进一步获得微电网的整体网络结构信息。

每一节点在特征计算中,都需要先统计影响该节点的邻域节点的注意力系数,以便于对各个邻域节点做出权重划分,使目标作用大的节点得到较高的注意力权重,而忽略目标影响较小的节点。对于中心节点特征向量 $\mathbf{h}_i$,利用注意力机制聚合节点的邻接节点本身特征向量 $\mathbf{h}_1, \dots, \mathbf{h}_j$ 。其中注意力权重系数 e_{ij} 代表 $\mathbf{h}_j$ 对 $\mathbf{h}_i$ 的影响程度, e_{ij} 的计算方式为

$$e_{ij} = \frac{\exp(\text{Relu}(a(\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_j)))}{\sum_{k=N_i} \exp(\text{Relu}(a(\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_k)))}$$

式中:Relu为激活函数; $\mathbf{W}$ 为可训练系数; $\parallel$ 为矩阵拼接符; a 为进行点积操作的参数。

最终,输入节点特征利用注意力系数 e_{ij} 进行加权求和后,得到 $\mathbf{h}_i$ 的新表达 $\mathbf{h}'_i$,计算方式为

$$\mathbf{h}'_i = \sigma \left(\sum_{j=N_i} e_{ij} \mathbf{W}\mathbf{h}_j \right) \quad (20)$$

采用两层图卷积网络对输入的特性进行提取,并全部采用Relu激活函数进行激活,最后通过全连接层输出最终动作。

2.3 GT-PPO算法

为更好地考虑电网内不同节点的相互作用对多能源微电网运行的影响,将区域内的节点及其交互关系构建一种动态图模型结构,并采用GAT对该非欧式结构的图模型数据进行特征提取,提取的特征以便进一步被PPO利用制定最优的运行策略。

微电网节点之间的相关性与其在电网中的位置有关。在径向微电网中,每个节点的电压都受到所有其他节点的影响,但影响随着距离的增加而减小。因此,智能体需要意识到网格中节点之间的关系,才能做出协同控制决策。为了利用电网节点之间的相关性,GT-PPO使用GAT对微电网内节点的相关性

进行建模,将电网状态信息投影为节点嵌入向量后输入PPO网络。

将多能源微电网运行优化策略问题转化为马尔可夫决策过程模型后,采用改进后的GT-PPO算法进行强化学习问题的求解。GT-PPO算法优化学习流程如图2所示。

在某一时段给定系统运行方式,通过数据采集得到环境状态,智能体根据环境状态给出动作,下一时段来临,环境根据动作和电网负荷、新能源发电等数据生成状态,后续交互均以此流程进行。同时智能体本身包括Actor网络和Critic网络,Actor网络和Critic网络共享较低层,有GNN和线性层组成。Actor网络输出动作,根据动作生成动作、状态和奖励的轨迹,轨迹输入Critic网络估计优势函数,更新两个网络的网络参数,最后求解出最优策略。

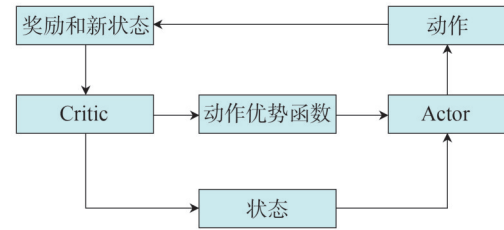


图2 算法优化学习流程

Fig.2 Algorithm optimization for learning process

3 案例分析

3.1 微电网场景设置

实验算例采用改进的微电网IEEE 33和IEEE 118节点系统进行仿真测试,验证所提方法的有效性和适应性。IEEE 33节点系统拓扑结构如图3所示,其中储能装置分别接在节点8,14,24,30上,风力发电机组和光伏发电机组分别接于节点5,10,16,20,26和32,35,36上。设节点1为平衡节点,其余均为PQ节点。每个网格都有一组场景,每个场景都指定了模拟中的变化,例如每个时间步的电源供应和需求。

3.2 算法及算例设置

分别采用GT-PPO算法、PPO算法、DDQN算法对微电网优化运行方案进行求解。在算法结构设置上,PPO算法和GT-PPO算法网络结构基本相同,隐藏层为128个神经元的3层全连接层;输出层设置4

个神经元作为动作输出数量。设置训练时更新迭代次数为10次,在训练过程中的具体参数设置如表1所示。

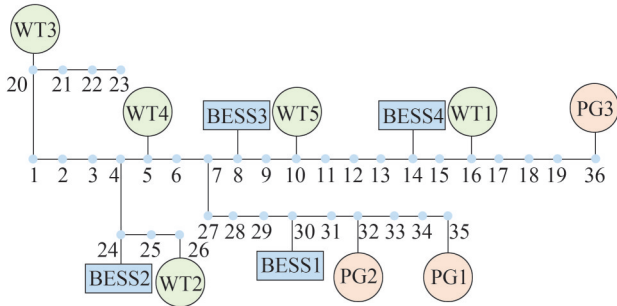


图3 改进的微电网 IEEE 33 节点系统

Fig.3 Diagram of the improved IEEE 33 node system for distribution network

表1 训练参数设置

Table 1 Training parameters setting

参数	参数值
训练集数 N	400
策略轨迹缓存大小 D	3 000
更新次数 I	10
小批量数据容量 B	200
奖励折扣因子 γ	0.99

3.3 算例结果及分析

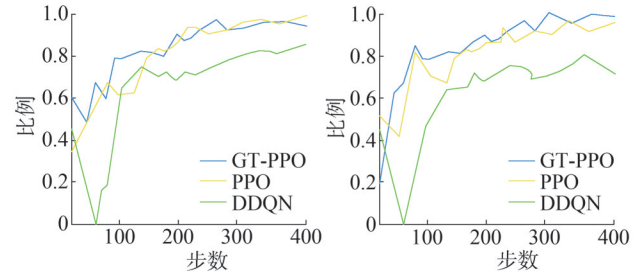
在实验中,我们使用可控比率、功率损耗和奖励函数得分3个指标来评估算法的性能。

可控比率是指计算在每个事件期间所有总线的电压都处于控制之下的时间步长的比率。功率损耗指的是计算每个事件中每个时间步长所有总线的总功率损耗的平均值。奖励函数得分取决于使用该算法所生成策略下,每一时间步下累计奖励的大小。

本文的目标是找到高可控速率和低功率损耗的算法。在训练过程中,设定总回合数 $N=1\ 000$ 。训练期间的测试结果由中位数给出。每20集进行一次测试,随机选取10集进行评估。

GT-PPO算法、PPO算法、DDQN算法在训练期间电压可控比率结果如图4所示,功率损耗如图5所示。结果表明,本文提出的GT-PPO算法在两种规模的电网上,性能均优于PPO算法和DDQN算法,在

电压可控比率更高的基础上,微电网运行的功率损耗更低。

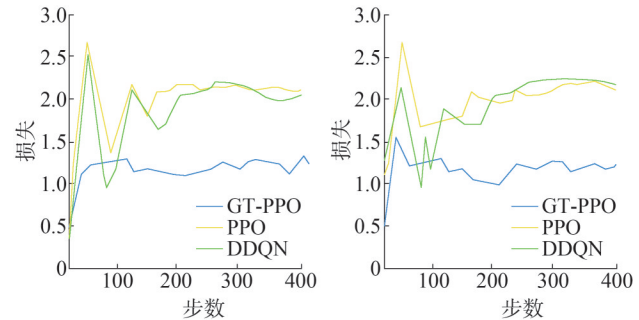


(a) IEEE 33

(b) IEEE 118

图4 训练期间(电压)可控比率

Fig.4 Controllable ratios (voltage) during training



(a) IEEE 33

(b) IEEE 118

图5 训练期间功率损耗

Fig.5 Power loss during training

在验证场景上评估推出的总平均缩放分数在 $[-100, 100]$ 范围内缩放,将无操作代理的返回值缩放并转换为0,表明该代理在安全和电力效率方面能够比无操作代理更好地管理电网。

在 IEEE 33 规模网络验证结果如图6所示, PPO、GT-PPO 和 DDQN 算法均能取得较好的效果,其中未改进的 PPO 算法的效果明显差于 DDQN 算法。使用改进后的 GT-PPO 算法效果得到了显著提升,评估得分高于 DDQN 算法。为了进一步验证算法的普适性,在改进后的 IEEE 118 节点电网上做进一步的拓展实验,结果如图7所示。在更大规模的 IEEE 118 网络中, PPO 和 DDQN 算法没能取得理想的效果,且由于动作空间和状态空间的扩大,算法变得不稳定, GT-PPO 算法效果相较于无代理更好,得分更高。

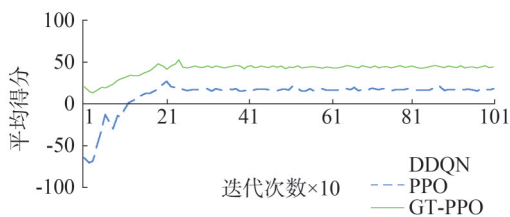


图6 IEEE 33验证集评估结果

Fig.6 IEEE 33 verification set evaluation results

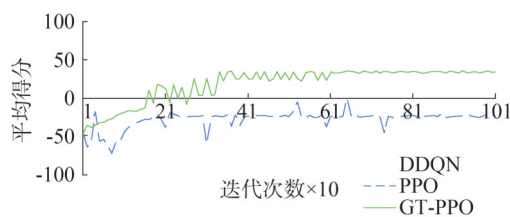


图7 IEEE 118验证集评估结果

Fig.7 IEEE 118 verification set evaluation results

本次实验使用全新的测试数据对训练完成后的模型进行测试。在测试时增加了MIQP算法对测试数据进行求解,MIQP算法不需要训练,没有训练数据。强化学习算法在改进后的IEEE 33节点网络和IEEE 118节点网络的测试得分如表2所示,可知所提GT-PPO算法在两个网络上得分均优于DDQN算法和PPO算法,其中所有算法在IEEE 33网络上均有很好的结果。但由于电网规模的增加,GT-PPO、DDQN和PPO算法得分分别为28.76、-4.75和8.32。

表2 算法得分结果

Table 2 Algorithm score result

强化学习算法	IEEE 33	IEEE 118
GT-PPO	68.56	28.76
DDQN	43.93	-4.75
PPO	45.87	8.32

可控比率和功率损耗的测试结果如表3和表4所示。测试结果表明,GT-PPO算法在维持电压稳定和减少功率损耗方面拥有更好的表现,在同样规模的电网下,相对于PPO算法和DDQN算法电压可控率更高且功率损耗更少。传统的MIQP方法在小规模电网上拥有较好的表现,略差于改进后的GT-PPO算法,在IEEE 118节点网络上,GT-PPO算法拥有更

好的适应性,电压可控比率和功率损耗结果均明显优于MIQP算法。

表3 IEEE 33电网测试结果

Table 3 IEEE 33 power grid test results

强化学习算法	可控比率/%	功率损耗/MW
GT-PPO	95.7	1.43
DDQN	93.4	2.21
PPO	82.1	2.15
MIQP	94.3	1.38

表4 IEEE 118规模电网测试结果

Table 4 IEEE 118 power grid test results

强化学习算法	可控比率/%	功率损耗/MW
GT-PPO	92.1	1.08
DDQN	87.6	2.53
PPO	72.1	2.19
MIQP	81.5	2.37

算例分析结果表明,提出的GT-PPO算法在微电网优化策略任务上可以取得较好的结果,且能够更加适应较大规模的电网。

4 结束语

针对微电网中源网荷储多因素的互动问题,考虑微电网运行的现状,研究储能参与微电网运行的能量互动机制以及与微电网内光伏、风机、用户负荷等进行协同调控的模式与策略。基于强化学习框架,通过将微电网运行问题模式建模为马尔可夫决策过程问题,然后以最小化微电网电压波动和运行损耗为目的构建微电网策略优化模型。设计一种GT-PPO算法,来学习并获得微电网最佳运行策略,实现微电网多能源微电网日常运行的电压波动减小和运行损耗降低。

算例分析表明,GT-PPO算法能够有效求解多能源微电网优化运行问题。在算法层面,GT-PPO算法继承了PPO算法的优点,具有更好的适应性能,对超参敏感度较低,同时结合图自注意力网络,充分利用微电网的互联结构,较之前的DDQN算法效果更好。合理优化微电网运行策略对多能源微电网稳定

运行有重要意义。

参考文献

- [1] 孙浩锋.“源网荷储”协调的主动配电网运行优化分析[D].郑州:郑州大学,2022.
- [2] 周磊,倪佳俊.智能化配电网运行方式的优化分析[J].集成电路应用,2022,39(6):272-273.
ZHOU Lei, NI Jiajun. Analysis of operation model optimization of intelligent distribution network[J]. Application of IC, 2022, 39(6): 272-273.
- [3] SANG J, SUN H, KOU L. Deep reinforcement learning microgrid optimization strategy considering priority flexible demand side[J]. Sensors: Basel, 2022, 22(6): 2256.
- [4] MASTERS C L. Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines [J]. Power Engineering Journal, 2002, 16(1): 5-12.
- [5] SENJYU T, MIYAZATO Y, YONA A, et al. Optimal distribution voltage control and coordination with distributed generation [J]. IEEE Transactions on Power Delivery, 2008, 23(2): 1236-1242.
- [6] AGALGAONKAR Y P, PAL B C, JABR R A. Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators [J]. IEEE Transactions on Power Systems, 2013, 29(1): 182-192.
- [7] VOVOS P N, KIPRAKIS A E, WALLACE A R, et al. Centralized and distributed voltage control: Impact on distributed generation penetration [J]. IEEE Transactions on Power Systems, 2007, 22(1): 476-483.
- [8] JAHANGIRI P, ALIPRANTIS D C. Distributed Volt/VAr control by PV inverters [J]. IEEE Transactions on Power Systems, 2013, 28(3): 3429-3439.
- [9] SCHIFFER J, SEEL T, RAISCH J, et al. Voltage stability and reactive power sharing in inverter-based microgrids with consensus-based distributed voltage control [J]. IEEE Transactions on Control Systems Technology, 2016, 24(1): 96-109.
- [10] SUN H B, GUO Q L, QI J J, et al. Review of challenges and research opportunities for voltage control in smart grids [J]. IEEE Transactions on Power Systems, 2019, 34(4): 2790-2801.
- [11] SINGHAL A, AJJARAPU V, FULLER J, et al. Real-time local volt/var control under external disturbances with high PV penetration [J]. IEEE Transactions on Smart Grid, 2019, 10(4): 3849-3859.
- [12] 刘一欣,郭力,王成山.微电网两阶段鲁棒优化经济调度方法[J].中国电机工程学报,2018,38(14):4013-4022.
LIU Yixin, GUO Li, WANG Chengshan. Economic dispatch of microgrid based on two stage robust optimization [J]. Proceedings of the CSEE, 2018, 38(14): 4013-4022.
- [13] 戴涛,孙伟军.采用模糊控制策略与主动阻尼控制技术的微电网能量管理技术[J].电测与仪表,2019,56(24):92-98.
DAI Tao, SUN Weijun. The power management method of micro-grid with fuzzy control strategy and active damping control technology [J]. Electrical Measurement & Instrumentation, 2019, 56(24): 92-98.
- [14] 张沛,朱驻军,谢桦.基于深度强化学习近端策略优化的电网无功优化方法[J].电网技术,2023,47(2):562-572.
ZHANG Pei, ZHU Zhujun, XIE Hua. Reactive power optimization based on proximal policy optimization of deep reinforcement learning [J]. Power System Technology, 2023, 47(2): 562-572.
- [15] 赵鹏杰,吴俊勇,王焱,等.基于深度强化学习的微电网优化运行策略[J].电力自动化设备,2022,42(11):9-16.
ZHAO Pengjie, WU Junyong, WANG Yi, et al. Optimal operation strategy of microgrid based on deep reinforcement learning [J]. Electric Power Automation Equipment, 2022, 42(11): 9-16.
- [16] YOON D, HONG S, LEE B J, et al. Winning the l2rpn challenge: Power grid management via semi-markov afterstate actor-critic [C]//International Conference on Learning Representations. 2021.
- [17] 刘俊峰,陈剑龙,王晓生,等.基于深度强化学习的微能源网能量管理与优化策略研究[J].电网技术,2020,44(10):3794-3803.
LIU Junfeng, CHEN Jianlong, WANG Xiaosheng, et al. Energy management and optimization of multi-energy grid based on deep reinforcement learning [J]. Power System Technology, 2020, 44(10): 3794-3803.
- [18] 熊正勇,苗虹,曾成碧.基于智能软开关与储能系统联合的有源配电网运行优化[J].电测与仪表,2020,57(13):33-39.
XIONG Zhengyong, MIAO Hong, ZENG Chengbi. Operation optimization of active distribution network based on smart soft open point and energy storage system [J]. Electrical Measurement & Instrumentation, 2020, 57(13): 33-39.
- [19] 李彦君,裴玮,肖浩,等.基于深度学习的微网需求响应特性封装与配电网优化运行[J].电力系统自动化,2021,45(10):157-165.
LI Yanjun, PEI Wei, XIAO Hao, et al. Deep learning based characteristic packaging of demand response for microgrids and optimal operation of distribution network. Automation of Electric Power Systems, 2021, 45(10): 157-165.
- [20] 刘宇,张聪,李涛.强化学习A3C算法在电梯调度中的建模及应用[J].计算机工程与设计,2022,43(1):196-202.
LIU Yu, ZHANG Cong, LI Tao. Modeling and application of reinforcement learning A3C in elevator scheduling algorithm [J]. Computer Engineering and Design, 2022, 43(1): 196-202.
- [21] WU G L, FANG W Q, WANG J, et al. Dyna-PPO reinforcement learning with Gaussian process for the continuous action decision-making in autonomous driving [J]. Applied Intelligence, 2023, 53(13): 16893-16907.
- [22] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms [EB/OL]. 2017: 1707.06347. <https://arxiv.org/abs/1707.06347v2>.

- [23] 王振宇,许静,胡文博,等.不确定性环境下园区风光储互动运行的PPO强化学习策略[J].电力需求侧管理,2022,24(5):44-50.
WANG Zhenyu, XU Jing, HU Wenbo, et al. PPO-based reinforcement learning strategy for interactive operation of wind-photovoltaic-storage in an uncertain environment[J].Power Demand Side Management,2022,24(5):44-50
- [24] 孙黎霞,白景涛,周照宇,等.基于双向长短期记忆网络的电力系统暂态稳定评估[J].电力系统自动化,2020,44(13):64-72.
SUN Lixia, BAI Jingtao, ZHOU Zhaoyu, et al. Transient stability assessment of power system based on bidirectional Long-short-term memory network [J]. Automation of Electric Power Systems, 2020,44 (13):64-72
- [25] 潮铸,段秦尉,钱峰,等.基于深度分布式强化学习的电力系统分散协调调度方法[J].电子器件,2022,45(4):947-953.
CHAO Zhu, DUAN Qinwei, QIAN Feng, et al. Decentralized coordinated dispatching method for power system based on deep distributed reinforcement learning[J].Chinese Journal of Electrom Devices,2022,45 (4):947-953.
- [26] 严梓铭,徐岩.结合深度强化学习与领域知识的电力系统拓扑结构优化[J].电力系统自动化,2022,46(1):60-68.
YAN Ziming, XU Yan. Topology optimization of power systems combining deep reinforcement learning and domain knowledge [J]. Automation of Electric Power Systems,2022,46 (1):60-68.

收稿日期:2023-04-12

修回日期:2023-10-16

作者简介:

李子凯(1990),男,工程师,从事电力系统优化、AI技术在电力系统的应用研究;

杨波(1970),男,高级工程师,从事电力系统自动化技术、微电网技术研究;

周忠堂(1980),男,经济师,从事电力大数据分析、新能源并网运行研究;

张健(1968),男,硕士,正高级经济师,从事AI技术在电力系统的应用研究;

徐明珠(1976),男,高级工程师,从事微电网与主电网协同优化运行和保护控制研究。

(责任编辑 姜婷婷)

(上接第11页)

- ZHENG Huankun, SUN Yaobin, WEI Kaiguo. Robust optimization and operation of hybrid AC/DC distribution grid interconnected by multiple flexible substations [J / OL]. Journal of North China Electric Power University (Natural Science Edition): 1-10 [2023-09-18]. <https://kns.cnki.net/kcms/detail/13.1212.TM.20230707.0854.002.html>.
- [17] 刘练,马小龙,潘万宝,等.基于精英策略布谷鸟算法的主动配电网经济优化调度[J].山东电力技术,2023,50(5):14-20.
LIU Lian, MA Xiaolong, PAN Wanbao, et al. Economic optimal dispatch of active distribution network based on elite strategy cuckoo algorithm [J]. Shandong Electric Power, 2023, 50 (5): 14-20.
- [18] 鹿优,鹿存鹏,徐伟,等.含抽水蓄能电站的多能互补微网系统设计与研究[J].山东电力技术,2023,50(5):34-40.
LU You, LU Cunpeng, XU Wei, et al. Design and research of multi-energy complementary microgrid system with pumped storage power station[J].Shandong Electric Power,2023,50(5):34-40.
- [19] 郭秉涛,谢宝昌,蔡旭.基于人工神经网络的海上风场双馈风机功率分配优化策略[J].电气自动化,2020,42(3):34-37.
- GUO Bingtao, XIE Baochang, CAI Xu. Power distribution optimization strategy for offshore wind power plant DFIG based on artificial neural network [J].Electrical Automation, 2020,42 (3): 34-37.
- [20] 孙辉,沈钟浩,周玮,等.电动汽车群响应的主动配电网阻塞调度研究[J].中国电机工程学报,2017,37(19):5549-5559.
SUN Hui, SHEN Zhonghao, ZHOU Wei, et al. Congestion dispatch research of active distribution network with electric vehicle group response[J].Proceedings of the CSEE,2017,37(19):5549-5559.

收稿日期:2023-09-20

修回日期:2023-11-06

作者简介:

江建(1973),男,硕士,高级工程师,主要研究方向为电力企业经营管理、电网规划建设、人力资源、安全生产等领域;

张树森(1982),男,硕士,高级工程师,主要研究方向为电网调控、生产经营、安全生产等领域;

徐峰亮(1977),男,高级工程师,主要研究方向为电力企业安全生产、信息通信、科技创新、储能、微电网等领域。

(责任编辑 车永强)