



(21) 申请号 202410410491.0

(22) 申请日 2024.04.07

(65) 同一申请的已公布的文献号

申请公布号 CN 118380997 A

(43) 申请公布日 2024.07.23

(73) 专利权人 国网湖北省电力有限公司电力科学研究院

地址 430077 湖北省武汉市洪山区徐东大街227号

(72) 发明人 杨增力 文博 田亦鎏 黎恒烜
周悦 张侃君 陈永昕 滕捷
史逸川 修连成(74) 专利代理机构 武汉楚天专利事务所 42113
专利代理师 孔敏

(51) Int.Cl.

H02J 3/00 (2006.01)

G06N 3/092 (2023.01)

G06Q 50/06 (2024.01)

H02J 3/24 (2006.01)

(56) 对比文件

CN 112668235 A, 2021.04.16

CN 116203839 A, 2023.06.02

审查员 曹玮

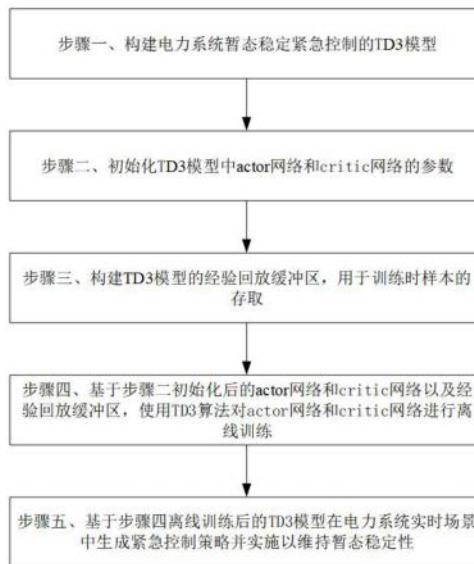
权利要求书4页 说明书8页 附图4页

(54) 发明名称

一种基于TD3的电力系统暂态稳定紧急控制策略生成方法

(57) 摘要

本发明提供一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,包括如下步骤:构建电力系统暂态稳定紧急控制的TD3模型;初始化TD3模型中actor网络和critic网络的参数;构建TD3模型的经验回放缓冲区;基于初始化后的actor网络和critic网络以及经验回放缓冲区,使用TD3算法对actor网络和critic网络进行离线训练;基于离线训练后的TD3模型在电力系统实时场景中生成紧急控制策略并实施以维持暂态稳定性。本发明实现了电力系统在面临暂态稳定性挑战时的紧急控制,通过训练actor网络和critic网络,能够生成有效的控制策略,提高电力系统在突发事件下的暂态稳定性。



1.一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,其特征在于,包括如下步骤:

步骤一、构建电力系统暂态稳定紧急控制的TD3模型;

步骤二、初始化TD3模型中actor网络和critic网络的参数;

步骤三、构建TD3模型的经验回放缓冲区,用于训练时样本的存取;

步骤四、基于步骤二初始化后的actor网络和critic网络以及步骤三构建的经验回放缓冲区,使用TD3算法对actor网络和critic网络进行离线训练;

步骤五、基于步骤四离线训练后的TD3模型在电力系统实时场景中生成紧急控制策略并实施以维持暂态稳定性;

所述步骤一具体包括:

步骤1.1:定义电力系统状态集合 $s_t = \{\delta_{Gi}(t), V_{Nj}(t)\}$, 其中 $\delta_{Gi}(t)$ 表示t时间的第i台发电机的功角, $V_{Nj}(t)$ 表示t时间的第j个节点的电压幅值, s_t 作为神经网络的输入层输入;

步骤1.2:定义电力系统输出动作集合 $a_t = \{k_{G1}, k_{G2}, \dots, k_{Gm}, k_{L1}, k_{L2}, \dots, k_{Ll}\}$, 其中 k_{Gi} 表示第i台发电机的切除系数, $i=1, 2, \dots, m$; k_{Lj} 表示第j个负荷的切除系数, $j=1, 2, \dots, l$, 切除系数 k 表示对当前发电机/负荷的切除量占当前发电机/负荷额定容量的比例, $0 \leq k \leq 1$;

发电机的切除控制按照整台切除来设置,即:

$$k_{Gi} = \begin{cases} 0, & \text{不进行切除措施} \\ 1, & \text{进行切除措施} \end{cases};$$

负荷的切除控制按照一定间隔 $p\%$ 来设置,即:

$$k_{Lj} = \{0, 1, 2, \dots, n\} \times p\%, 0 < p \leq 100, n = \frac{1}{p\%};$$

步骤1.3:定义电力系统暂态稳定判断标准 TSI :

$$TSI = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|};$$

其中 $|\Delta\delta_{\max}|$ 表示步骤3.1中任意两台发电机中最大相对功角差, $TSI > 0$ 时,系统稳定,反之系统失稳;

步骤1.4:定义即时奖励函数 R_t :

$$R_t = R_{state} + R_{others};$$

$$R_{state} = \begin{cases} r_{stable}, & TSI > 0 \\ r_{unstable}, & TSI \leq 0 \end{cases};$$

$$R_{others} = r_c + r_v + r_s;$$

其中:

R_{state} 表示控制后系统状态的奖励值, r_{stable} 表示系统稳定时候的 R_{state} 值, 取正数, $r_{unstable}$ 表示系统不稳定时候的 R_{state} 值, 取负数;

R_{others} 表示控制后其他奖励值;

r_c 表示控制成本奖励值, 根据切除量的加权和给予惩罚, 反映切机和切负荷的成本:

$$r_c = \sum_{i=1}^m c_G P_{Gi} + \sum_{j=1}^l c_L P_{Lj};$$

$$P_{Gi} = P_{ratedGi} k_{Gi}, P_{Lj} = P_{ratedLj} k_{Lj};$$

P_{Gi} 表示第 i 台发电机的切除量, $P_{ratedGi}$ 表示第 i 台发电机的额定容量; P_{Lj} 表示第 j 个负荷的切除量, $P_{ratedLj}$ 表示第 j 个负荷的额定容量, m 为发电机数量, l 为负荷数量;

r_v 表示电压约束奖励值, 根据超出上下限的大小给予惩罚, 以限制控制后各节点电压:

$$r_v = c_v \sum_{n=1}^N r_n;$$

$$r_n = \begin{cases} V_n - V_n^{\max}, & V_n > V_n^{\max} \\ V_n^{\min} - V_n, & V_n < V_n^{\min} \\ 0, & V_n^{\min} < V_n < V_n^{\max} \end{cases};$$

c_v 表示电压越限惩罚系数, r_n 表示节点 n 电压的越限值, V_n 表示第 n 个节点的电压值, $V_n^{\max}$ 和 $V_n^{\min}$ 分别表示第 n 个节点的电压幅值的上下限;

r_s 表示控制步数奖励, 根据控制步数的总次数给予惩罚, 以最少的动作次数完成控制目标:

$$r_s = c_s N_{step};$$

c_s 表示控制步数惩罚系数, N_{step} 表示控制次数;

所述步骤五具体包括:

步骤5.1: 获取当前电力系统的实时状态 s_t , 包括发电机功角 $\delta_{Gi}(t)$ 、节点电压 $V_{Nj}(t)$;

步骤5.2: 输入当前状态到已经离线训练好的TD3模型的 actor 网络中, 生成相应的紧急控制动作 a_t ;

步骤5.3: 将生成的紧急控制动作应用到电力系统中, 实施紧急控制策略;

步骤5.4: 实时监测实施后的电力系统状态响应, 包括 TSI 、电压暂态稳定性指标, 并利用响应信息更新 actor 网络中的实时状态表示, 以保持网络的准确性和泛化能力;

步骤5.5: 如果系统响应符合预期, 继续维持当前控制策略; 如果系统响应不理想, 将状态信息添加到经验回放缓冲区中, 供后续离线训练和改进控制策略。

2. 如权利要求1所述的基于TD3的电力系统暂态稳定紧急控制策略生成方法, 其特征在

于:所述步骤二具体包括:

步骤2.1:在 actor 网络中,初始化包括状态输入层、多个全连接隐藏层和输出层的神经网络参数,其中输入数据是电力系统状态集合 s_t ,输出数据是电力系统输出动作集合 a_t ,输出层的节点数等于电力系统的可行动作空间维度 D_{action} ,以确保生成合理的控制动作, D_{action} 为电力系统输出动作集合 a 的长度;

步骤2.2:在 critic 网络中,初始化包括状态输入层、动作输入层、多个全连接隐藏层和输出层的神经网络参数,其中输入数据为电力系统状态和输出动作集合 $\{s_t, a_t\}$,输出层的节点数为1,输出数据为对应状态和输出动作组合的Q值 $Q(s_t, a_t)$:

$$Q(s_t, a_t) = R_t + \gamma \cdot \max_{a_{t+1}} Q(s_{t+1}, a_{t+1});$$

其中:

$Q(s_t, a_t)$ 表示状态 s_t 下采取动作 a_t 的Q值;

R_t 表示状态 s_t 下采取动作 a_t 的即时奖励;

γ 是折扣因子,用于平衡当前奖励和未来奖励的重要性;

s_{t+1} 表示采取动作 a_t 后转移到的下一个状态;

a_{t+1} 表示 s_{t+1} 状态下选择的动作;

$\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$ 表示在下一个状态 s_{t+1} 选择最优动作的Q值;

步骤2.3:设定离线训练的迭代次数、每次迭代中的训练步数和收敛标准。

3.如权利要求2所述的基于TD3的电力系统暂态稳定紧急控制策略生成方法,其特征在于:所述步骤三具体包括:

步骤3.1:采集电力系统仿真数据,包括系统状态 s_t 和 s_{t+1} ,控制动作 a_t ,并根据数据计算对应奖励 R_t ;

步骤3.2 设定经验回放缓冲区的存储容量,并按照时间顺序将采集到的数据以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储到缓冲区中,构建经验回放缓冲区。

4.如权利要求3所述的基于TD3的电力系统暂态稳定紧急控制策略生成方法,其特征在于:所述步骤四具体包括:

步骤4.1:随机从经验回放缓冲区抽样批量的经验数据作为训练样本;

步骤4.2:对训练样本使用 actor 网络生成当前状态下 s_t 的控制动作 a_t ,通过仿真获得执行控制动作后系统状态 s_{t+1} ,计算得到奖励值 R_t ,并通过 critic 网络评估该动作的Q值;

步骤4.3:使用双Q学习方法,计算当前状态下的目标值 Q_t ,并计算 critic 网络输出的Q值与目标值 Q_t 的均方误差;

步骤4.4:利用均方误差最小化的优化算法更新 critic 网络的参数;

步骤4.5:计算 actor 网络的策略梯度;

步骤4.6:利用梯度上升法,更新 actor 网络的参数,以提高生成控制动作的质量;

步骤4.7:以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储于经验缓冲区作为经验数据;

步骤4.8 重复以上步骤,直到达到预定的训练步数或达到收敛标准完成离线训练。

一种基于TD3的电力系统暂态稳定紧急控制策略生成方法

技术领域

[0001] 本发明涉及电力系统控制领域,具体是一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,用于提高电力系统在面临暂态失稳情况下的稳定性和鲁棒性。

背景技术

[0002] 电力系统的暂态稳定性是指电力系统在遭受外部扰动后,能够在一定时间内恢复到新的平衡状态的能力。在面临突发故障或负荷波动时,电力系统可能面临暂态失稳的风险,导致电力系统崩溃或发生停电。因此,开发高效的暂态稳定控制策略对于确保电力系统的可靠运行至关重要。

[0003] 传统的电力系统控制方法主要依赖于经验规则和线性控制器,这些方法在复杂的非线性和时变电力系统中表现出限制。随着深度强化学习技术的发展,利用强化学习算法来设计电力系统暂态稳定控制策略成为一种备受关注的方法。然而,现有的方法往往面临收敛速度慢、稳定性差以及在实际电力系统中泛化能力不足等挑战。在实际电力系统中,由于系统的非线性和不确定性,传统控制方法难以提供足够鲁棒性的控制策略。因此,需要一种能够充分考虑电力系统的动态特性、非线性特性和多变因素的电力系统暂态稳定紧急控制策略生成方法。

发明内容

[0004] 本发明的目的在于提供一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,通过深度强化学习的框架,结合离线经验回放缓冲区和双Q学习的思想,实现了电力系统在面临暂态稳定性挑战时的紧急控制,通过训练actor网络和critic网络,该方法能够生成有效的控制策略,提高电力系统在突发事件下的暂态稳定性。

[0005] 一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,包括如下步骤:

[0006] 步骤一、构建电力系统暂态稳定紧急控制的TD3模型;

[0007] 步骤二、初始化TD3模型中actor网络和critic网络的参数;

[0008] 步骤三、构建TD3模型的经验回放缓冲区,用于训练时样本的存取;

[0009] 步骤四、基于步骤二初始化后的actor网络和critic网络以及步骤三构建的经验回放缓冲区,使用TD3算法对actor网络和critic网络进行离线训练;

[0010] 步骤五、基于步骤四离线训练后的TD3模型在电力系统实时场景中生成紧急控制策略并实施以维持暂态稳定性。

[0011] 进一步的,所述步骤一具体包括:

[0012] 步骤1.1:定义电力系统状态集合 $s_t = \{\delta_{Gi}(t), V_{Nj}(t)\}$,其中 $\delta_{Gi}(t)$ 表示t时间的第i台发电机的功角, $V_{Nj}(t)$ 表示t时间的第j个节点的电压幅值, s_t 作为神经网络的输入层输入;

[0013] 步骤1.2:定义电力系统输出动作集合 $a_t = \{k_{G1}, k_{G2}, \dots, k_{Gm}, k_{L1}, k_{L1}, \dots, k_{Ll}\}$,其中 k_{Gi} 表示第i台发电机的切除系数, $i = 1, 2, \dots, m$; k_{Lj} 表示第j个负荷的切除系数, $j = 1,$

2, ..., 1, 切除系数k表示对当前发电机/负荷的切除量占当前发电机/负荷额定容量的比例, $0 \leq k \leq 1$;

[0014] 发电机的切除控制按照整台切除来设置, 即:

$$[0015] \quad k_{Gi} = \begin{cases} 0, & \text{不进行切除措施} \\ 1, & \text{进行切除措施} \end{cases}$$

[0016] 负荷的切除控制按照一定间隔p%来设置, 即:

$$[0017] \quad k_{Lj} = \{0, 1, 2, \dots, n\} \times p\%, 0 < p \leq 100, n = \frac{1}{p\%}$$

[0018] 步骤1.3: 定义电力系统暂态稳定判断标准TSI:

$$[0019] \quad TSI = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|}$$

[0020] 其中 $|\Delta\delta_{\max}|$ 表示步骤3.1中任意两台发电机中最大相对功角差, TSI > 0时,

[0021] 系统稳定, 反之系统失稳;

[0022] 步骤1.4: 定义即时奖励函数 R_t :

$$[0023] \quad R_t = R_{\text{state}} + R_{\text{others}}$$

$$[0024] \quad R_{\text{state}} = \begin{cases} r_{\text{stable}}, & TSI > 0 \\ r_{\text{unstable}}, & TSI \leq 0 \end{cases}$$

$$[0025] \quad R_{\text{others}} = r_c + r_v + r_s$$

[0026] 其中:

[0027] R_{state} 表示控制后系统状态的奖励值, r_{stable} 表示系统稳定时候的 R_{state} 值, 取正数, r_{unstable} 表示系统不稳定时候的 R_{state} 值, 取负数;

[0028] R_{others} 表示控制后其他奖励值;

[0029] r_c 表示控制成本奖励值, 根据切除量的加权和给予惩罚, 反映切机和切负荷的成本:

$$[0030] \quad r_c = \sum_{i=1}^m c_G P_{Gi} + \sum_{j=1}^l c_L P_{Lj}$$

$$[0031] \quad P_{Gi} = P_{\text{rated}Gi} k_{Gi}, P_{Lj} = P_{\text{rated}Lj} k_{Lj}$$

[0032] P_{Gi} 表示第i台发电机的切除量, $P_{\text{rated}Gi}$ 表示第i台发电机的额定容量; P_{Lj} 表示第i个负荷的切除量, $P_{\text{rated}Lj}$ 表示第i个负荷的额定容量。m为发电机数量, l为负荷数量;

[0033] r_v 表示电压约束奖励值, 根据超出上下限的大小给予惩罚, 以限制控制后各节点电压:

$$[0034] \quad r_v = c_v \sum_{n=1}^N r_n$$

$$[0035] \quad r_n = \begin{cases} V_n - V_n^{\max}, & V_n > V_n^{\max} \\ V_n^{\min} - V_n, & V_n < V_n^{\min} \\ 0, & V_n^{\min} < V_n < V_n^{\max} \end{cases}$$

[0036] c_v 表示电压越限惩罚系数, r_n 表示节点n电压的越限值, V_n 表示第n个节点的电压值, $V_n^{\max}$ 和 $V_n^{\min}$ 分别表示第n个节点的电压幅值的上下限;

[0037] r_s 表示控制步数奖励,根据控制步数的总次数给予惩罚,以最少的动作次

[0038] 数完成控制目标:

$$[0039] \quad r_s = c_s N_{\text{step}}$$

[0040] c_s 表示控制步数惩罚系数, N_{step} 表示控制次数。

[0041] 进一步的,所述步骤二具体包括:

[0042] 步骤2.1:在actor网络中,初始化包括状态输入层、多个全连接隐藏层和输出层的神经网络参数,其中输入数据是电力系统状态集合 s_t ,输出数据是电力系统输出动作集合 a_t ,输出层的节点数等于电力系统的可行动作空间维度 D_{action} ,以确保生成合理的控制动作, D_{action} 为电力系统输出动作集合 α 的长度;

[0043] 步骤2.2:在critic网络中,初始化包括状态输入层、动作输入层、多个全连接隐藏层和输出层的神经网络参数,其中输入数据为电力系统状态和输出动作集合 $\{s_t, a_t\}$,输出层的节点数为1,输出数据为对应状态和输出动作组合的Q值 $Q(s_t, a_t)$:

$$[0044] \quad Q(s_t, a_t) = R_t + \gamma \cdot \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$$

[0045] 其中:

[0046] $Q(s_t, a_t)$ 表示状态 s_t 下采取动作 a_t 的Q值;

[0047] R_t 表示状态 s_t 下采取动作 a_t 的即时奖励;

[0048] γ 是折扣因子,用于平衡当前奖励和未来奖励的重要性;

[0049] s_{t+1} 表示采取动作 a_t 后转移到的下一个状态;

[0050] a_{t+1} 表示 s_{t+1} 状态下选择的动作;

[0051] $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$ 表示在下一个状态 s_{t+1} 选择最优动作的Q值;

[0052] 步骤2.3:设定离线训练的迭代次数、每次迭代中的训练步数和收敛标准。

[0053] 进一步的,所述步骤三具体包括:

[0054] 步骤3.1:采集电力系统仿真数据,包括系统状态 s_t 和 s_{t+1} ,控制动作 a_t ,并根据数据计算对应奖励 R_t ;

[0055] 步骤3.2设定经验回放缓冲区的存储容量,并按照时间顺序将采集到的数据以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储到缓冲区中,构建经验回放缓冲区。

[0056] 进一步的,所述步骤四具体包括:

[0057] 步骤4.1:随机从经验回放缓冲区抽样批量的经验数据作为训练样本;

[0058] 步骤4.2:对训练样本使用actor网络生成当前状态下 s_t 的控制动作 a_t ,通过仿真获得执行控制动作后系统状态 s_{t+1} ,计算得到奖励值 R_t ,并通过critic网络评估该动作的Q值;

[0059] 步骤4.3:使用双Q学习方法,计算当前状态下的目标值 Q_t ,并计算critic网络输出

的Q值与目标值 Q_t 的均方误差;

[0060] 步骤4.4:利用均方误差最小化的优化算法更新critic网络的参数;

[0061] 步骤4.5:计算actor网络的策略梯度;

[0062] 步骤4.6:利用梯度上升法,更新actor网络的参数,以提高生成控制动作的质量;

[0063] 步骤4.7:以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储于经验缓冲区作为经验数据;

[0064] 步骤4.8重复以上步骤,直到达到预定的训练步数或达到收敛标准完成离线训练。

[0065] 进一步的,所述步骤五具体包括:

[0066] 步骤5.1:获取当前电力系统的实时状态 s_t ,包括发电机功角 $\delta_{Gi}(t)$ 、节点电压 $V_{Nj}(t)$;

[0067] 步骤5.2:输入当前状态到已经离线训练好的TD3模型的actor网络中,生成相应的紧急控制动作 a_t ;

[0068] 步骤5.3:将生成的紧急控制动作应用到电力系统中,实施紧急控制策略;

[0069] 步骤5.4:实时监测实施后的电力系统状态响应,包括TSI、电压等暂态稳定性指标,并利用响应信息更新actor网络中的实时状态表示,以保持网络的准确性和泛化能力;

[0070] 步骤5.5:如果系统响应符合预期,继续维持当前控制策略;如果系统响应不理想,将状态信息添加到经验回放缓冲区中,供后续离线训练和改进控制策略。

[0071] 本发明具有强大的场景适应性和卓越的实用性,通过引入深度强化学习算法TD3,能够有效地学习和优化电力系统的暂态稳定控制策略。电力系统属于一种高维的状态空间,在面临暂态稳定问题时的控制策略属于连续动作空间,相比于其它深度强化学习算法,TD3能够综合考虑多个时间步的信息,更全面地学习电力系统的动态演化,因而能够更灵活地处理高维状态空间,并且有效地应对连续动作空间的问题,使其能够在实时控制场景中迅速应对电力系统的动态变化,同时,其采用的经验回放缓冲区和双Q学习技术也有效提高了算法的稳定性和效率,进一步提升了其性能。因此,本发明为电力系统暂态稳定性的提升提供了一种创新的、鲁棒的解决方案。

附图说明

[0072] 图1是本发明实施例一种基于TD3的电力系统暂态稳定紧急控制策略生成方法的流程示意图;

[0073] 图2是本发明实施例WEPRI-36节点系统架构图;

[0074] 图3是采用本发明实施例紧急控制策略前的功角曲线;

[0075] 图4是采用本发明实施例紧急控制策略前的仿真图;

[0076] 图5是采用本发明实施例紧急控制策略后的功角曲线;

[0077] 图6是采用本发明实施例紧急控制策略后的仿真图。

具体实施方式

[0078] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明的一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0079] 请参阅图1,本发明实施例提供一种基于TD3的电力系统暂态稳定紧急控制策略生成方法,包括如下步骤:

[0080] 步骤一、构建电力系统暂态稳定紧急控制的TD3模型;

[0081] 所述步骤一具体包括:

[0082] 步骤1.1:定义电力系统状态集合 $s_t = \{\delta_{Gi}(t), V_{Nj}(t)\}$,其中 $\delta_{Gi}(t)$ 表示t时间的第i台发电机的功角, $V_{Nj}(t)$ 表示t时间的第j个节点的电压幅值, s_t 作为神经网络的输入层输入;

[0083] 步骤1.2:定义电力系统输出动作集合 $a_t = \{k_{G1}, k_{G2}, \dots, k_{Gm}, k_{L1}, k_{L1}, \dots, k_{Ll}\}$,其中 k_{Gi} 表示第i台发电机的切除系数, $i = 1, 2, \dots, m$; k_{Lj} 表示第j个负荷的切除系数, $j = 1, 2, \dots, l$,切除系数k表示对当前发电机/负荷的切除量占当前发电机/负荷额定容量的比例, $0 \leq k \leq 1$;

[0084] 发电机的切除控制按照整台切除来设置,即:

$$[0085] \quad k_{Gi} = \begin{cases} 0, & \text{不进行切除措施} \\ 1, & \text{进行切除措施} \end{cases}$$

[0086] 负荷的切除控制按照一定间隔p%来设置,即:

$$[0087] \quad k_{Lj} = \{0, 1, 2, \dots, n\} \times p\%, 0 < p \leq 100, n = \frac{1}{p\%}$$

[0088] 步骤1.3:定义电力系统暂态稳定判断标准TSI:

$$[0089] \quad TSI = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|}$$

[0090] 其中 $|\Delta\delta_{\max}|$ 表示步骤3.1中任意两台发电机中最大相对功角差,TSI>0时,

[0091] 系统稳定,反之系统失稳;

[0092] 步骤1.4:定义即时奖励函数 R_t :

$$[0093] \quad R_t = R_{\text{state}} + R_{\text{others}}$$

$$[0094] \quad R_{\text{state}} = \begin{cases} r_{\text{stable}}, & TSI > 0 \\ r_{\text{unstable}}, & TSI \leq 0 \end{cases}$$

$$[0095] \quad R_{\text{others}} = r_c + r_v + r_s$$

[0096] 其中:

[0097] R_{state} 表示控制后系统状态的奖励值, r_{stable} 表示系统稳定时候的 R_{state} 值,取正数, r_{unstable} 表示系统不稳定时候的 R_{state} 值,取负数;

[0098] R_{others} 表示控制后其他奖励值;

[0099] r_c 表示控制成本奖励值,根据切除量的加权和给予惩罚,反映切机和切负荷的成本:

$$[0100] \quad r_c = \sum_{i=1}^m c_G P_{Gi} + \sum_{j=1}^l c_L P_{Lj}$$

$$[0101] \quad P_{Gi} = P_{\text{ratedGi}} k_{Gi}, P_{Lj} = P_{\text{ratedLj}} k_{Lj}$$

[0102] P_{Gi} 表示第i台发电机的切除量, P_{ratedGi} 表示第i台发电机的额定容量; P_{Lj} 表示第j个

负荷的切除量, $P_{\text{rated},j}$ 表示第 i 个负荷的额定容量。 m 为发电机数量, l 为负荷数量;

[0103] r_v 表示电压约束奖励值, 根据超出上下限的大小给予惩罚, 以限制控制后各节点电压:

$$[0104] \quad r_v = c_v \sum_{n=1}^N r_n$$

$$[0105] \quad r_n = \begin{cases} V_n - V_n^{\max}, & V_n > V_n^{\max} \\ V_n^{\min} - V_n, & V_n < V_n^{\min} \\ 0, & V_n^{\min} < V_n < V_n^{\max} \end{cases}$$

[0106] c_v 表示电压越限惩罚系数, r_n 表示节点 n 电压的越限值, V_n 表示第 n 个节点的电压值, $V_n^{\max}$ 和 $V_n^{\min}$ 分别表示第 n 个节点的电压幅值的上下限;

[0107] r_s 表示控制步数奖励, 根据控制步数的总次数给予惩罚, 以最少的动作次数完成控制目标:

$$[0108] \quad r_s = c_s N_{\text{step}}$$

[0109] c_s 表示控制步数惩罚系数, N_{step} 表示控制次数。

[0110] 步骤1.1-1.4构建了电力系统暂态稳定紧急控制的TD3模型框架, 通过框架的设定, 使该方法能够在具有动态特性、非线性特性和多变因素的电力系统遇到暂态稳定故障问题时提供有效的控制策略, 并且显著提高TD3离线训练的效率。

[0111] 步骤二、初始化TD3模型中actor网络和critic网络的参数;

[0112] 所述步骤二具体包括:

[0113] 步骤2.1: 在actor网络中, 初始化包括状态输入层、多个全连接隐藏层和输出层的神经网络参数, 其中输入数据是电力系统状态集合 s_t , 输出数据是电力系统输出动作集合 a_t , 输出层的节点数等于电力系统的可行动作空间维度 D_{action} , 以确保生成合理的控制动作, D_{action} 为电力系统输出动作集合 α 的长度;

[0114] 步骤2.2: 在critic网络中, 初始化包括状态输入层、动作输入层、多个全连接隐藏层和输出层的神经网络参数, 其中输入数据为电力系统状态和输出动作集合 $\{s_t, a_t\}$, 输出层的节点数为1, 输出数据为对应状态和输出动作组合的Q值 $Q(s_t, a_t)$:

$$[0115] \quad Q(s_t, a_t) = R_t + \gamma \cdot \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$$

[0116] 其中:

[0117] $Q(s_t, a_t)$ 表示状态 s_t 下采取动作 a_t 的Q值;

[0118] R_t 表示状态 s_t 下采取动作 a_t 的即时奖励;

[0119] γ 是折扣因子, 用于平衡当前奖励和未来奖励的重要性;

[0120] s_{t+1} 表示采取动作 a_t 后转移到的下一个状态;

[0121] a_{t+1} 表示 s_{t+1} 状态下选择的动作;

[0122] $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$ 表示在下一个状态 s_{t+1} 选择最优动作的Q值;

[0123] 步骤2.3: 设定离线训练的迭代次数、每次迭代中的训练步数和收敛标准。

[0124] 步骤2.1-2.3设置了actor和critic神经网络的参数, 通过参数的设定, 能够帮助

TD3模型实现更好的训练效果。

[0125] 步骤三、构建TD3模型的经验回放缓冲区,用于训练时样本的存取;

[0126] 所述步骤三具体包括:

[0127] 步骤3.1:采集电力系统仿真数据,包括系统状态 s_t 和 s_{t+1} ,控制动作 a_t ,并根据数据计算对应奖励 R_t ;

[0128] 步骤3.2设定经验回放缓冲区的存储容量,并按照时间顺序将采集到的数据以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储到缓冲区中,构建经验回放缓冲区。

[0129] 步骤3.1-3.2构建了TD3模型的经验回放缓冲区,用于训练时样本的存取,通过经验回放缓冲区的构建,能够降低数据的相关性并且提高模型的训练效率。

[0130] 步骤四、基于步骤二初始化后的actor网络和critic网络以及步骤三构建的经验回放缓冲区,使用TD3算法对actor网络和critic网络进行离线训练;

[0131] 所述步骤四具体包括:

[0132] 步骤4.1:随机从经验回放缓冲区抽样批量的经验数据作为训练样本;

[0133] 步骤4.2:对训练样本使用actor网络生成当前状态下 s_t 的控制动作 a_t ,通过仿真获得执行控制动作后系统状态 s_{t+1} ,计算得到奖励值 R_t ,并通过critic网络评估该动作的Q值;

[0134] 步骤4.3:使用双Q学习方法,计算当前状态下的目标值 Q_t ,并计算critic网络输出的Q值与目标值 Q_t 的均方误差;

[0135] 步骤4.4:利用均方误差最小化的优化算法更新critic网络的参数;

[0136] 步骤4.5:计算actor网络的策略梯度;

[0137] 步骤4.6:利用梯度上升法,更新actor网络的参数,以提高生成控制动作的质量;

[0138] 步骤4.7:以 $e_t = \{s_t, a_t, R_t, s_{t+1}\}$ 的格式存储于经验缓冲区作为经验数据;

[0139] 步骤4.8重复以上步骤,直到达到预定的训练步数或达到收敛标准完成离线训练。

[0140] 步骤4.1-4.8完成了TD3模型的离线训练过程。完成离线训练后的TD3模型将能够根据输入状态 s_t 生成对应的输出动作 a_t ,此时的输出动作 a_t 即为针对当前输入状态 s_t 生成的紧急控制动作。

[0141] 步骤五、基于步骤四离线训练后的TD3模型在电力系统实时场景中生成紧急控制策略并实施以维持暂态稳定性。

[0142] 所述步骤五具体包括:

[0143] 步骤5.1:获取当前电力系统的实时状态 s_t ,包括发电机功角 $\delta_{Gi}(t)$ 、节点电压 $V_{Nj}(t)$;

[0144] 步骤5.2:输入当前状态到已经离线训练好的TD3模型的actor网络中,生成相应的紧急控制动作 a_t ;

[0145] 步骤5.3:将生成的紧急控制动作应用到电力系统中,实施紧急控制策略;

[0146] 步骤5.4:实时监测实施后的电力系统状态响应,包括TSI、电压等暂态稳定性指标,并利用响应信息更新actor网络中的实时状态表示,以保持网络的准确性和泛化能力;

[0147] 步骤5.5:如果系统响应符合预期,继续维持当前控制策略;如果系统响应不理想,将状态信息添加到经验回放缓冲区中,供后续离线训练和改进控制策略。

[0148] 本发明适用于电力系统稳定控制技术领域,用于恢复电力系统在发生严重故障后的暂态稳定性,针对以上问题,本实施例通过PSASP7.41采用基频为50Hz的WEPRI-36节点电

力系统进行仿真,用于说明紧急控制策略生成方法效果。该系统含有8台发电机,10个负荷,系统架构如图2所示。TD3模型及训练基于PyTorch 1.12.1搭建,采用Python3.7语言进行编程。

[0149] 模型状态空间包括8台发电机的功角、36个节点的母线电压,故状态空间维度数为44。设置分段切除负荷的比例 $p\%$ 为10%,输出动作空间包括8个切机动作和100个切负荷动作,动作空间维度数为108。Actor网络和Critic网络的隐藏层层数均为3层,每层包含的神经元个数分别为50,25和50。Critic网络学习率为0.0005,Actor网络学习率为0.0001,经验回放池容量设置为 1×10^5 ,经验回放训练批次为64,折旧因子为0.90,软更新系数为0.001。

[0150] 初始时刻系统处于稳定状态,设置离线学习故障场景为输电线AC34的50%处发生三相短路故障,故障持续时间为0.2s,设置学习的目标为故障切除后5s内保持暂态稳定。

[0151] 本实施例采用暂态稳定指数(Transient Stability Index,TSI)作为稳定判据:

$$[0152] \quad TSI = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|}$$

[0153] 式中 $|\Delta\delta_{\max}|$ 为仿真时长内任意两台发电机相对功角差的最大值。当 $TSI > 0$ 时,系统状态为稳定,反之,判定为失稳状态。

[0154] 当故障场景为输电线AC34的50%处时,0.02s发生三相短路故障,0.22s故障切除。在不实施任何紧急控制策略的情况下,通过暂稳仿真得到结果如图3和图4所示,从图3可以看出,4、7、8号发电机的功角曲线与其他发电机功角曲线失去同调性,从图4可以看出,发电机最大功角差在1s时已经越过了360度基准线以下, $TSI < 0$,说明系统处于失稳状态。

[0155] 根据TD3模型可得此场景下的紧急控制策略如表1所示,在仿真中应用该控制策略,得到仿真结果如图5和图6所示,从图5可以看出,在切除7号发电机和母线23负荷的10%之后,发电机的功角曲线恢复同调性,同时从图6可以看出,发电机最大功角差一直保持在360度基准线以下, $TSI > 0$,说明系统在执行了紧急控制策略之后恢复了稳定性,可见该方法的有效性。

[0156] 表1 AC34三相短路故障下紧急控制策略结果

[0157]	控制步数	控制动作	切除量/MW	TSI	奖励值 R
	1	切除发电机 G7	286	<0(不稳定)	-5.064
	2	切除母线 23 负荷 10%	26.28	>0(稳定)	+122.073
	控制结果			稳定	117.009

[0158] 本发明利用TD3算法中的离线学习机制,通过不断在仿真环境中采样、经验回放、更新网络参数的方式,优化控制策略,以适应电力系统不同工况下的紧急控制需求。同时,采用目标网络的软更新机制,有效减缓网络更新的速度,提高控制策略的稳定性。本发明不仅提高了电力系统的暂态稳定性,还具备一定的泛化能力,能够适应不同电力系统结构和工况下的紧急控制问题。在实际应用中,本发明为电力系统提供了一种智能、自适应的暂态稳定紧急控制方案,为电力系统运行安全和可靠性提供了新的解决思路。

[0159] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何属于本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到的变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应该以权利要求的保护范围为准。

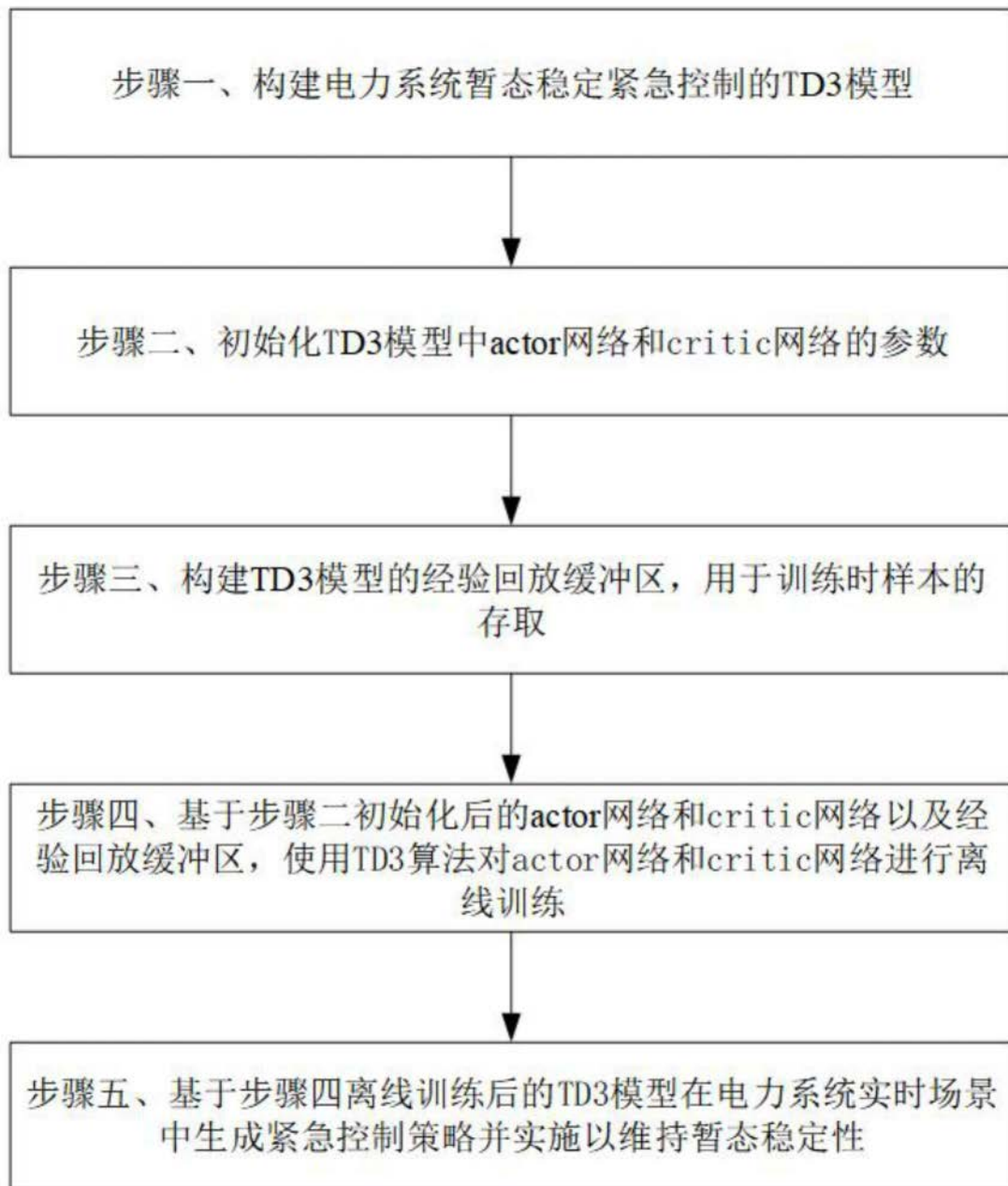


图1

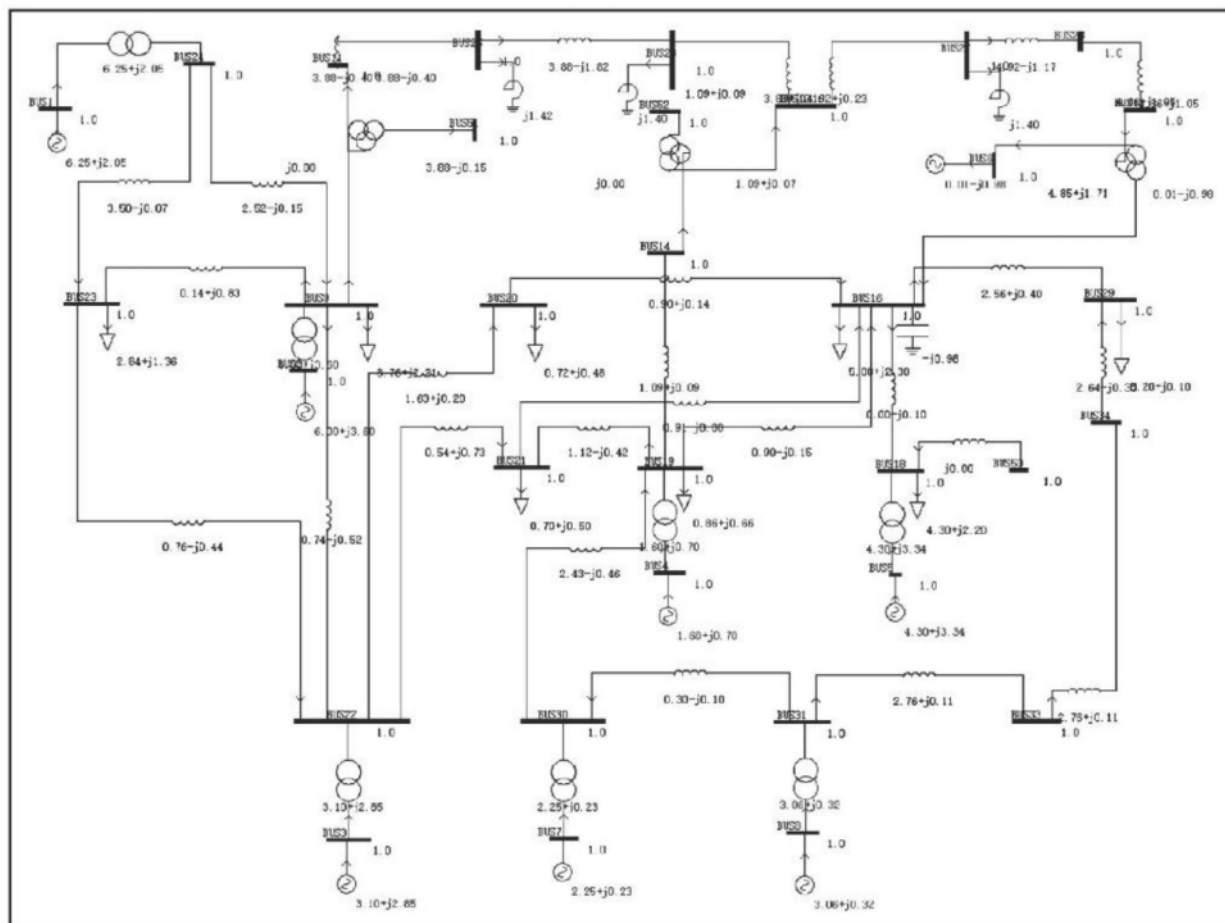


图2

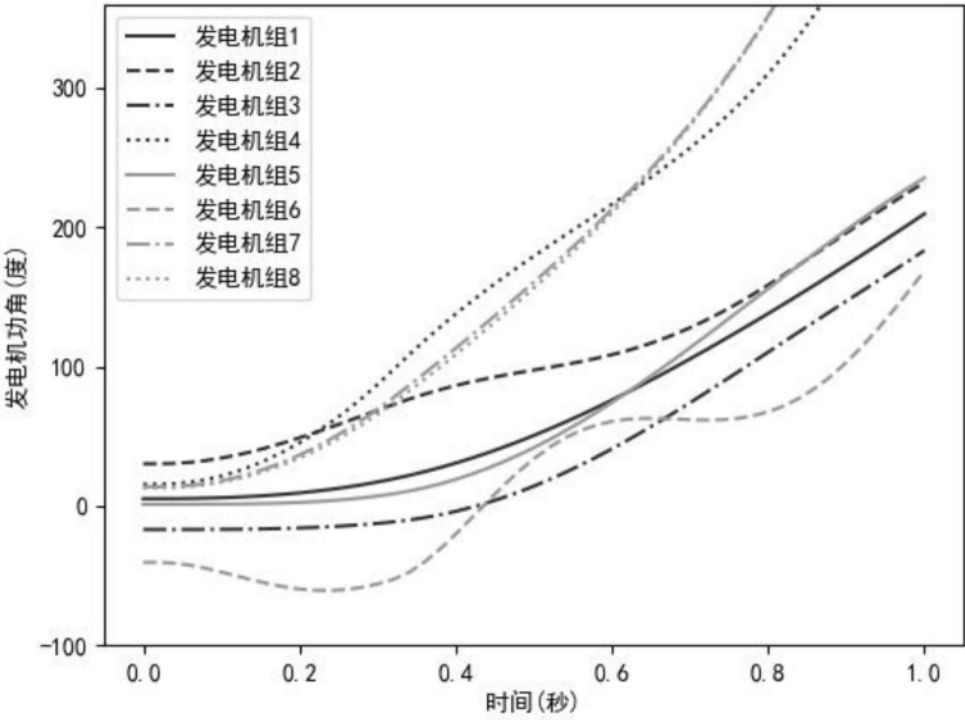


图3

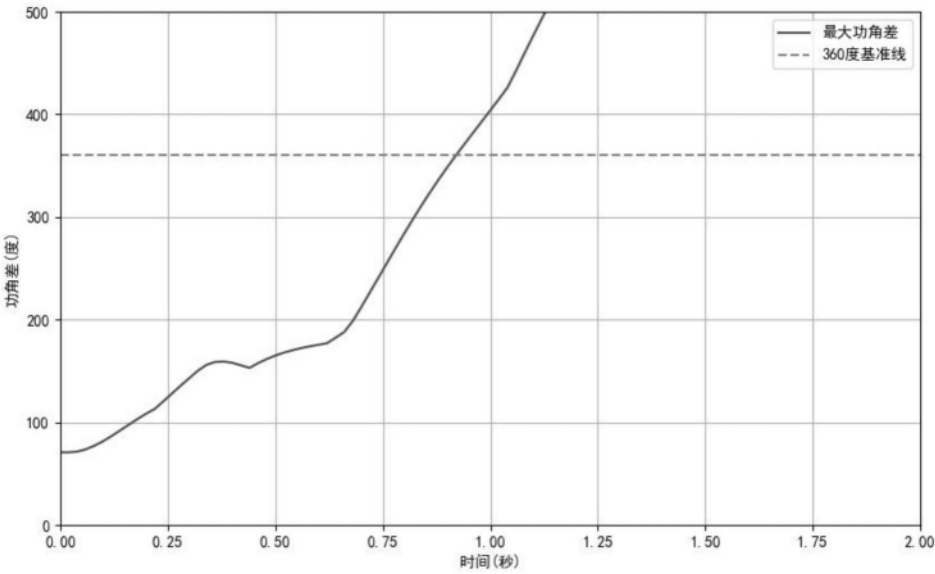


图4

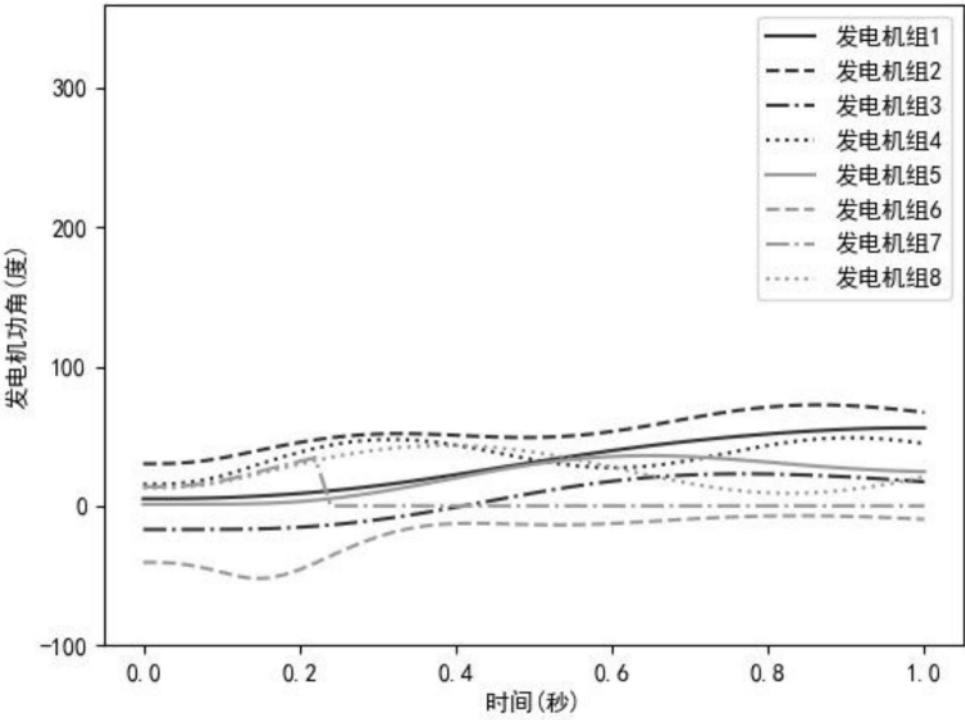


图5

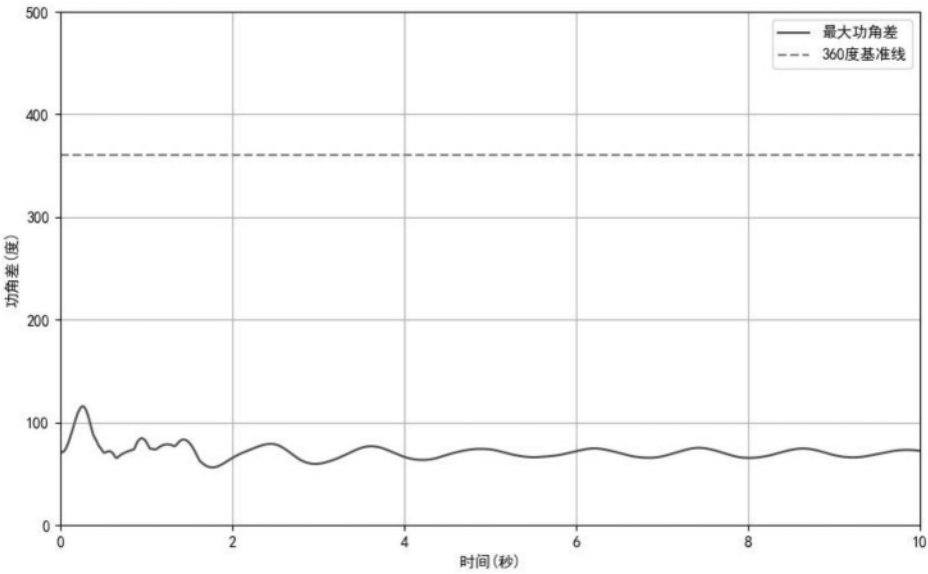


图6