

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 8 月 4 日 (04.08.2022)



(10) 国际公布号
WO 2022/160705 A1

- (51) 国际专利分类号:
G06Q 10/06 (2012.01) *G06Q 50/06* (2012.01) [CN/CN]; 中国北京市西城区西长安街 86 号, Beijing 100031 (CN)。
- (21) 国际申请号: PCT/CN2021/115751
- (22) 国际申请日: 2021 年 8 月 31 日 (31.08.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110107505.8 2021 年 1 月 26 日 (26.01.2021) CN
- (71) 申请人: 中国电力科学研究院有限公司 (CHINA ELECTRIC POWER RESEARCH INSTITUTE COMPANY LIMITED) [CN/CN]; 中国北京市海淀区清河小营东路 15 号, Beijing 100192 (CN)。 国网天津市电力公司 (STATE GRID TIANJIN ELECTRIC POWER COMPANY) [CN/CN]; 中国天津市河北区五经路 39 号, Tianjin 300010 (CN)。 国家电网有限公司 (STATE GRID CORPORATION OF CHINA)
- (72) 发明人: 乔骥 (QIAO, Ji); 中国北京市海淀区清河小营东路 15 号, Beijing 100192 (CN)。 王新迎 (WANG, Xinying); 中国北京市海淀区清河小营东路 15 号, Beijing 100192 (CN)。 蒲天骄 (PU, Tianjiao); 中国北京市海淀区清河小营东路 15 号, Beijing 100192 (CN)。
- (74) 代理人: 北京派特恩知识产权代理有限公司 (CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区海淀南路 21 号中关村知识产权大厦 B 座 2 层, Beijing 100080 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) **Title:** METHOD AND APPARATUS FOR CONSTRUCTING DISPATCHING MODEL OF INTEGRATED ENERGY SYSTEM, MEDIUM, AND ELECTRONIC DEVICE

(54) 发明名称: 综合能源系统调度模型构建方法、装置、介质及电子设备

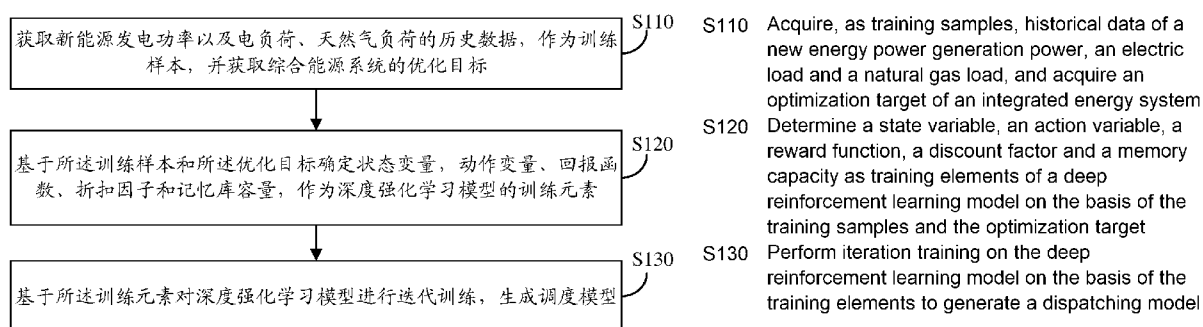


图 1

(57) **Abstract:** A dispatching method and apparatus for an integrated energy system, a medium and an electronic device. Said method comprises: collecting, as training samples, historical data of a new energy power generation power, an electric load and a natural gas load, and acquiring an optimization target of an integrated energy system (S110); determining a state variable, an action variable, a reward function, a discount factor and a memory capacity as training elements of a deep reinforcement learning model on the basis of the training samples and the optimization target (S120); and iterating the deep reinforcement learning model on the basis of the training elements to generate a dispatching model (S130).

(57) **摘要:** 一种综合能源系统的调度方法、装置、介质及电子设备。所述方法包括: 收集新能源发电功率以及电负荷、天然气负荷的历史数据, 作为训练样本, 并获取综合能源系统的优化目标 (S110); 基于所述训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子、记忆库容量, 作为深度强化学习模型的训练元素 (S120); 基于所述训练元素对深度强化学习模型进行迭代, 生成调度模型 (S130)。

LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

综合能源系统调度模型构建方法、装置、介质及电子设备

相关申请的交叉引用

本申请基于申请号为 202110107505.8、申请日为 2021 年 01 月 26 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的
5 全部内容在此以引入方式并入本申请。

技术领域

本申请实施例涉及能源互联网综合能源系统优化运行领域，尤其涉及一种综合能源系统调度模型构建方法、装置、介质及电子设备。

背景技术

10 随着科学技术的不断发展，当前世界范围内日益增长的能源需求与环境生态保护之间矛盾凸显，提高能源利用效率，寻求替代性新能源已成为世界各国能源变革的必经之路。灵活的燃气轮机组为平抑间歇性新能源出力提供了保障；对于电网难以消纳的新能源出力，可将其通过电转气（P2G，Power to Gas）技术转化为天然气/氢气大规模存储于天然气管道中。电力系
15 统与天然气系统逐步呈现深度耦合的趋势，通过构建充分协同的电-气互联综合能源系统，提高能源利用效率，以实现系统间的灵活运行与多能互补。

能源互联网综合能源系统的优化运行基础理论及求解问题涉及的控制参量众多，运行状态随机多变，各类状态变量高度耦合，往往为一个复杂的非线性、非凸优化问题。同时，能源互联网集中控制方式随着控制设备
20 的增多，计算量呈指数级增加。因此，采用传统计算方法难以解决能源互联网的优化运行问题。

发明内容

本申请实施例提供一种综合能源系统调度模型构建方法、装置、介质及电子设备，能够传统方法收敛速度慢、决策时间长、无法适应源荷随机性场景、需要大量简化假设等问题。

5 第一方面，本申请实施例提供了一种综合能源系统调度模型构建方法，所述方法包括：获取新能源发电功率以及电负荷、天然气负荷的历史数据，作为训练样本，并获取综合能源系统的优化目标；基于所述训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素；基于所述训练元素对深度强化学习模
10 型进行迭代训练，生成调度模型。

本申请实施例提供的综合能源系统调度模型构建方法，通过深度强化学习与环境的不断交互，更新神经网络参数，并自适应学习电-气耦合能源系统的源荷不确定性，无需引入机理简化假设；能够在线、实时地给出优化调度策略。生成连续控制策略，相比于传统离散型强化学习模型，降低
15 了动作离散化导致的计算误差。能够适用于新能源随机多场景下的优化调度策略，且具有良好的可扩展性。

在一些可选实施例中，所述优化目标包括：运行成本最小、碳排放量最小、新能源消纳率最高中的至少一种。

20 在一些可选实施例中，所述深度强化学习模型包括：行动器和评判器；所述基于所述训练元素对所述深度强化学习模型进行迭代训练，包括：

对构建的深度强化学习模型进行初始化；所述初始化包括设置训练迭代次数和每一次迭代的调度周期；

从调度周期内的第一个时刻开始，计算该时刻的综合能源系统的能流分布，得到系统环境的状态变量，并利用回报函数计算此刻系统的运行成本；
25 本；

将所述状态变量输入至深度强化学习模型的行动器中，输出调度策略作为动作变量；

所述调度策略叠加一个正态分布的噪声后，作用于综合能源系统，得到下一状态变量；

5 将所述状态变量、所述动作变量、运行成本、所述下一状态变量以及是否为终止状态的 5 元组，存入强化学习智能体的经验回放单元；并从强化学习智能体的经验回放单元中抽取预设数量的样本，基于抽取的样本对行动器和评价器的参数进行更新，持续迭代，直至迭代收敛稳定，得到训练好的深度强化学习模型。

10 在本实施例中，训练时采用深度学习对输入的综合能源系统的状态变量，输出调度策略作为动作变量，将调度策略作用于综合能源后继续生成下一个状态变量，此过程中与环境不断交互，更新网络参数，采用克服传统方法收敛速度慢、决策时间长、无法适应源荷随机性场景、需要大量简化假设等问题，通过深度神经网络自动、自适应地学习源荷概率分布，能够
15 在秒级内快速、在线地给出优化调度策略。

在一些可选实施例中，所述训练元素中：所述状态变量，表示各时刻系统的能流分布；所述动作变量，表示系统的控制变量，包括常规机组发电功率、燃气轮机功率、电转气功率、储能充放电功率中的至少之一；所述回报函数，表示每时刻系统的优化目标分类；所述记忆库容量，表示强
20 化学习智能体的记忆库容量，当智能体存储的 5 元组个数超过该容量值时，则随机删除记忆库中的某一个 5 元组，存储新的 5 元组。

在一些可选实施例中，所述方法还包括：获取未来调度周期内的新能源发电功率以及电负荷、天然气负荷需求，在所述未来调度周期内的每一时间步，向所述调度模型的行动器中输入系统的运行状态，生成各时刻的
25 调度策略。

在一些可选实施例中，所述行动器包括行动器在线策略网络及其目标策略网络；所述评判器包括评判器在线网络及其目标网络；所述基于抽取的样本对行动器和评价器的参数进行更新，包括：

设置折扣因子等价式；

- 5 将折扣因子等价式代入预设损失函数中计算，采用预设损失函数更新评判器在线网络 Q 的参数；

采用下式更新行动器的在线策略网络 μ ：

$$\nabla_{\theta^{\mu}} J = \frac{1}{m} \sum_n \nabla_a Q(s, a | \theta^Q) \big|_{s=s_n, a=\mu(s_n)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) \big|_{s_n}$$

- 其中，m 为随机抽取的样本数量， $\nabla_{\theta^{\mu}} \mu$ 表示对策略网络 μ 的网络参数
10 求偏导数， ∇_a 表示是对输入的动作 a 求偏导数； θ^{μ} 表示在线对策略网络 μ 的参数； θ^Q 为评判器在线网络 Q 的参数。

采用软更新方式，更新行动器目标网络 μ' 和评判器目标网络 Q' 。

在一些可选实施例中，所述预设损失函数采用如下公式：

$$L = \frac{1}{m} \sum_n (y_n - Q(s_n, a_n | \theta^Q))^2$$

- 15 其中， y_n 表示折扣因子等价式； $y_n = r_n + \gamma Q'(s_{n+1}, \mu'(s_{n+1} | \theta^{\mu'})) | \theta^Q$ ， $n = 1, 2 \dots m$ ； θ^Q 为评判器在线网络 Q 的参数； S_n 为当前状态变量， a_n 为调度策略。

第二方面，本申请实施例提供了一种综合能源系统的调度模型构建装置，所述装置包括：

- 20 训练样本收集模块，配置为获取新能源发电功率以及电负荷、天然气负荷的历史数据，作为训练样本，并获取综合能源系统的优化目标；

模型构建模块，配置为基于所述从训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素；

调度策略迭代模块，配置为基于所述训练元素对深度强化学习模型进

行迭代训练，生成调度模型。

第三方面，本申请实施例提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如本申请实施例所述的综合能源系统调度模型构建方法。

5 第四方面，本申请实施例提供了一种电子设备，包括存储器，处理器及存储在存储器上并可在处理器运行的计算机程序，所述处理器执行所述计算机程序时实现如本申请实施例所述的综合能源系统调度模型构建方法。

本申请实施例所提供的技术方案，收集新能源发电功率以及电负荷、
10 天然气负荷的历史数据，作为训练样本，并获取综合能源系统的优化目标；
基于所述训练样本和所述优化目标确定状态变量，动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素；基于所述训练元素对深度强化学习模型进行迭代训练，生成调度模型。该方法能够克服传统方法收敛速度慢、决策时间长、无法适应源荷随机性场景、需要大
15 量简化假设等问题，通过深度神经网络自动、自适应地学习源荷概率分布，训练好的模型，能够在秒级内快速、在线地给出优化调度策略。

附图说明

构成本申请的一部分的说明书附图用来提供对本发明的进一步理解，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不
20 当限定。在附图中：

图 1 是本申请实施例提供的综合能源系统的调度模型构建方法的流程图；

图 2 (a) 是本申请实施例提供的一种综合能源系统的调度模型中深度强化学习的示意图；

25 图 2 (b) 是本申请实施例提供的一种综合能源系统调度模型构建方法

的模型训练方法的流程图;

图 3 是本申请实施例提供的综合能源系统的调度装置的结构示意图;

图 4 是本申请实施例提供的一种电子设备的结构示意图。

具体实施方式

5 下面将参考附图并结合实施例来详细说明本申请。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

以下详细说明均是示例性的说明,旨在对本申请提供进一步的详细说明。除非另有指明,本申请所采用的所有技术术语与本申请所属领域的一般技术人员的通常理解的含义相同。本申请所使用的术语仅是为了描述具
10 体实施方式,而并非意图限制根据本申请的示例性实施方式。

随着人工智能技术的兴起和发展,强化学习在电力系统的优化控制中越来越受到重视。强化学习模型通过与环境的不断交互,实现经验的逐渐积累和策略的持续改进。尤其是深度神经网络与强化学习结合后的深度强化学习模型,具有更好的自适应学习能力和非凸优化、非线性问题的优化
15 决策能力,适合于解决综合能源系统的优化调度问题。基于此,本申请将深度强化学习模型应用于综合能源系统,以生成调度策略。

图 1 是本申请实施例提供的综合能源系统的调度方法的流程图,本实施例可适用于综合能源系统的调度的情况,该方法可以由本申请实施例所提供的综合能源系统调度模型构建装置执行,该装置可以由软件和/或硬件
20 的方式来实现,并可集成于电子设备中。

如图 1 所示,综合能源系统的调度模型构建方法包括:

S110、获取新能源发电功率以及电负荷、天然气负荷的历史数据,作为训练样本,并获取综合能源系统的优化目标。

具体的,收集新能源(光伏、风电等)发电功率以及电负荷、天然气
25 负荷的历史数据,存入数据库中,作为训练样本。

在本实施例中，可选的，所述优化目标包括：运行成本最小、碳排放量最少、新能源消纳率最高中的至少一种。

示例性的，所述获得综合能源系统的优化目标，例如可以是运行成本最小、碳排放量最小、新能源消纳率最高等优化目标中的至少一种目标。

5 S120、基于所述训练样本和所述优化目标确定状态变量，动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素。

其中，示例性的，状态变量，表示各时刻系统的能流分布；

动作变量，表示系统的控制变量，包括常规机组发电功率、燃气轮机功率、电转气功率、储能充放电功率中的至少之一；

10 回报函数，表示每时刻系统的优化目标分类；

记忆库容量，表示强化学习智能体的记忆库容量，当智能体存储的 5 元组个数超过该容量值时，则随机删除记忆库中的某一个 5 元组，存储新的 5 元组。

15 在一些可选实施例中，所述深度强化学习模型包括：行动器、在线策略网络及其目标策略网络、评判器、以及在线网络及其目标网络。

面向电-气耦合的综合能源系统，构建相应的深度强化学习模型，包括状态变量 S 、动作变量 A 、回报函数 R 、折扣因子 γ 和记忆库容量 M ；强化学习模型包含四个神经网络，分别为行动器（Actor）在线策略网络 μ 及其目标策略网络 μ' ，和评判器（Critic）在线网络 Q 及其目标网络 Q' 。

20 在本实施例中，可选的，所述行动器、所述在线策略网络及其目标策略网络、所述评判器、以及所述在线网络及其目标网络均采用全连接神经网络构建。

S130、基于所述训练元素对深度强化学习模型进行迭代训练，生成调度模型。

25 在一些可选实施例中，所述深度强化学习模型包括：行动器和评判器；

所述基于所述训练元素对所述深度强化学习模型进行迭代训练，包括：对构建的深度强化学习模型进行初始化；所述初始化包括设置训练迭代次数和每一次迭代的调度周期；从调度周期内的第一个时刻开始，计算该时刻的综合能源系统的能流分布，得到系统环境的状态变量，以及此刻系统的运行成本；将所述状态变量输入至深度强化学习模型的行动器中，输出调度策略作为动作变量；所述调度策略叠加一个正态分布的噪声后，作用于综合能源系统，得到下一状态变量；将所述状态变量、所述调度策略、所述下一状态变量、回报函数以及是否为终止状态的 5 元组，存入强化学习智能体的经验回放单元；并从强化学习智能体的经验回放单元中抽取预设数量的样本，对行动器和评价器进行更新，持续迭代，直至迭代收敛稳定，得到训练好的深度强化学习模型。

如图 2 (a) 所示，设总的训练迭代轮数为 N ，每一轮的调度周期为 T ，调度策略生成的时间间隔为 λ ，则每一轮调度周期共需生成 $p=T/\lambda$ 次调度策略。从调度周期内的第一个时刻 $t_n(n=0)$ 开始，计算该时刻的综合能源系统的能流分布，可得到系统环境的状态变量 S_n ，以及此刻系统的运行成本 R_n 。 S_n 输入至强化学习模型的行动器 μ 中，输出调度策略（即动作变量） A_n 。动作变量 A_n 作用于能源系统，得到系统的下一个状态 S_{n+1} 。

将当前状态变量 S_n 、动作变量 A_n 、下一状态变量 S_{n+1} 、奖励（或运行成本） R_n ，是否为终止状态 is_end 的 5 元组存入智能体的经验回放单元。

从智能体的经验回放单元中随机抽取 m 个样本，得到采样集合。

采用采样集合（即 m 个样本）更新行动器 μ 和评判器 Q 。

采用软更新方式，更新行动器目标网络 μ' 和评判器目标网络 Q' 。

可选的，在对行动器和评价器进行参数更新之后，所述方法还包括：持续迭代，直至迭代收敛稳定，得到深度强化学习模型。

上述过程持续进行，直至迭代收敛稳定，保存模型。此时整个调度周

期内的累积回报函数 R 达到最大。

在上述技术方案的基础上，可选的，在得到调度模型之后，所述方法还包括：获得未来调度周期内的新能源发电功率以及电负荷、天然气负荷需求，在所述未来调度周期内的每一时间步，向所述调度模型的行动器中
5 输入系统的运行状态，生成各时刻的调度策略。

采用训练好的模型进行未来调度策略生成。预测未来一个调度周期内的新能源（光伏、风电等）发电功率以及电负荷、天然气负荷需求，在调度周期内的每一时间步，向调度模型的行动器中输入系统的运行状态，逐个生成各时刻的调度策略。

10 图 2（b）是本申请实施例提供的一种综合能源系统调度模型构建方法的模型训练方法的流程示意图。

在模型构建以及模型训练之前，包括：

步骤 101，收集新能源（例如光伏、风电等）发电功率以及电负荷、天然气负荷的历史数据，存入数据库中，作为训练样本。示例性的，历史数
15 据的采样周期与调度周期一致，例如每一个小时生成一次调度策略，则训练样本中的各历史数据也为每小时一个数据点。

步骤 102，设置综合能源系统的优化目标，包括系统运行成本、碳排放量、新能源消纳率等目标中的至少一种。

步骤 2 为深度强化学习模型的构建过程，参见图 2（b）所示，具体包
20 括：

步骤 201，构建强化学习模型的各项元素，例如包括：状态变量 S （即各时刻系统的能流分布）；动作变量 A （即系统的控制变量，例如常规机组发电功率，燃气轮机功率，电转气（P2G）功率，储能充放电功率等）；回报函数 R （即每时刻系统的优化目标分类）；折扣因子 γ ；强化学习智能体的
25 记忆库容量 M 等等。

其中，在构建强化学习模型的各项元素之前，初始化行动-评判网络参数，清空经验回放单元 D。其中，行动-评判网络包括行动器和评判器；行动器包括行动器在线策略网络及其目标策略网络；所述评判器包括评判器在线网络及其目标网络。进一步，初始化电气耦合系统（或者也可称为综合能源系统）当前状态 S_n 。

步骤 202，强化学习模型包含四个神经网络，分别为行动器（Actor）在线策略网络 μ 及其目标策略网络 μ' ，和评判器（Critic）在线网络 Q 及其目标网络 Q' ，均采用全连接神经网络构建。在线策略网络 μ 及其目标策略网络 μ' 结构相同，网络参数分别为 θ^μ 和 $\theta^{\mu'}$ ，输入层维度为状态变量 S 的个数，输出为动作的个数；评判器在线网络 Q 及其目标网络 Q' 结构相同，网络参数分别为 θ^Q 和 $\theta^{Q'}$ ，输出维度为 1。

步骤 3 具体包括：

步骤 301，设总的训练迭代轮数为 N ，每一轮的调度周期为 T ，调度策略生成的时间间隔为 λ ，则每一轮调度周期共需生成 $p=T/\lambda$ 次调度策略。

步骤 302，从调度周期内的第一个时刻 $t_n(n=0)$ 开始，计算该时刻的综合能源系统（或电气耦合系统）的能流分布，可得到系统环境的当前状态变量 S_n ，以及此刻系统的运行成本 R_n 。

步骤 303，在行动器中基于状态 S 得到动作 A ，即将状态变量 S_n 输入至强化学习模型的行动器 μ 中，输出作为动作变量的调度策略 A_n 。

步骤 304，将动作变量作用于仿真环境，即动作变量 A_n 叠加一个正态分布（正态分布的方程随迭代的进行逐渐减小）的噪声后，作用于能源系统，得到系统的下一个状态变量 S_{n+1} 。

步骤 4，将当前状态变量 S_n 、动作变量 A_n 、下一个状态变量 S_{n+1} 、运行成本 R_n 、是否为终止状态 is_end 的 5 元组 $\{S_n, A_n, S_{n+1}, R_n, is_end\}$ 存入强化学习智能体的经验回放单元 D。这里经验回放单元 D 中的训练样本是否充

足，若充足，执行步骤 5；若不充足，则重新获得当前状态变量 S_n 。

步骤 5，从经验回放单元中随机抽取 m 个样本， m 为正整数。

步骤 6，是对评判器 Q 和行动器 μ 的参数更新过程，包括使用均方差损失函数，通过梯度反向传播更新评判器在线策略网络的参数；更新行动器
5 常在线策略网络的参数，进而采用软更新的方式更新行动器目标网络和评判器目标网络的参数。具体如下：

步骤 601，设置折扣因子等价式；如下公式 1 所示：

$$y_n = r_n + \gamma Q'(s_{n+1}, \mu'(s_{n+1} | \theta^{\mu'}) | \theta^Q), \quad n = 1, 2 \dots m \quad (\text{公式 1})$$

然后，将折扣因子等价式代入如下公式 2 的损失函数中，采用如下损
10 失函数更新评判器在线网络的参数：

$$L = \frac{1}{m} \sum_n (y_n - Q(s_n, a_n | \theta^Q))^2 \quad (\text{公式 2})$$

其中， y_n 表示折扣因子等价式； θ^Q 为评判器在线网络 Q 的参数； $\theta^{Q'}$ 为目标网络 Q' 的参数； S_n 为当前状态变量， a_n 为调度策略。

步骤 602 采用下式更新行动器的在线策略网络 μ 的参数：

$$\nabla_{\theta^{\mu}} J = \frac{1}{m} \sum_n \nabla_a Q(s, a | \theta^Q) |_{s=s_n, a=\mu(s_n)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) |_{s_n} \quad (\text{公式 3})$$

m 为随机抽取的样本数量， $\nabla_{\theta^{\mu}} \mu$ 表示对策略网络 μ 的网络参数求偏导数， ∇_a 表示是对输入的动作 a （如动作变量）求偏导数； θ^{μ} 表示在线对策略网络 μ 的参数； θ^Q 为评判器在线网络 Q 的参数。

步骤 7，采用软更新方式，更新行动器目标网络 μ' 和评判器目标网络 Q'
20 的参数，即

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \quad (\text{公式 4})$$

$$\theta^{\mu'} \leftarrow \tau \theta^{\mu} + (1 - \tau) \theta^{\mu'} \quad (\text{公式 5})$$

其中， τ 为更新系数，可设置为 0.01。

步骤 8, 上述步骤 3-步骤 7 的过程持续进行, 直至迭代收敛稳定, 即判定当前为终止状态, 保存训练好的深度强化学习模型。此时整个调度周期内的累积回报函数 R 达到最大。

步骤 9, 采用训练好的深度强化学习模型 (即调度模型) 生成未来调度周期的调度策略。预测未来一个调度周期内的新能源 (光伏、风电等) 发电功率以及电负荷、天然气负荷需求, 在调度周期内的每一时间步, 向强化学习模型的行动器中输入系统的运行状态, 逐个生成各时刻的调度策略。

本申请实施例不依赖于具体的物理模型, 因此也不需要进行大量的简化假设, 而是直接自适应地学习源荷随机性分布, 避免了因随机性简化导致的计算误差。训练后的强化学习模型能够针对不同的随机调度场景, 在线快速给出调度策略, 无需重新进行模型训练, 策略生成速度更快。

本申请实施例属于无模型的优化方法, 具有较好地可扩展性。相比于传统方法, 生成的策略为连续控制策略, 降低了因为控制动作离散化导致的计算误差。

本申请实施例通过深度强化学习与环境的不断交互, 更新神经网络参数, 并自适应学习电-气耦合能源系统的源荷不确定性, 无需引入机理简化假设; 能够在线、实时地给出优化调度策略。生成连续控制策略, 相比于传统离散型强化学习模型, 降低了动作离散化导致的计算误差。该方法能够适用于新能源随机多场景下的优化调度策略, 且具有良好的可扩展性。

图 3 是本申请实施例提供的综合能源系统调度模型构建装置的结构示意图。如图 3 所示, 所述装置包括:

训练样本收集模块 310, 配置为获取新能源发电功率以及电负荷、天然气负荷的历史数据, 作为训练样本, 并获取综合能源系统的优化目标;

模型构建模块 320, 配置为构建深度强化学习模型, 基于所述训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子和记忆库

容量，作为深度强化学习模型的训练元素；

调度策略迭代模块 330，配置为基于所述训练元素对深度强化学习模型进行迭代训练，生成调度模型。

在本申请的一些可选实施例中，所述优化目标包括：运行成本最小、碳排放量最小、新能源消纳率最高中的至少一种。

在本申请的一些可选实施例中，所述深度强化学习模型包括：行动器和评判器；所述调度策略迭代模块 330，配置为对构建的深度强化学习模型进行初始化；所述初始化包括设置训练迭代次数和每一次迭代的调度周期；从调度周期内的第一个时刻开始，计算该时刻的综合能源系统的能流分布，得到系统环境的状态变量，并利用回报函数计算此刻系统的运行成本；将所述状态变量输入至深度强化学习模型的行动器中，输出调度策略作为动作变量；所述调度策略叠加一个正态分布的噪声后，作用于综合能源系统，得到下一状态变量；将所述状态变量、所述动作变量、运行成本、所述下一状态变量以及是否为终止状态的 5 元组，存入强化学习智能体的经验回放单元；并从强化学习智能体的经验回放单元中抽取预设数量的样本，基于抽取的样本对行动器和评价器的参数进行更新，持续迭代，直至迭代收敛稳定，得到训练好的深度强化学习模型。

在本申请的一些可选实施例中，所述行动器包括行动器在线策略网络及其目标策略网络；所述评判器包括评判器在线网络及其目标网络；所述调度策略迭代模块 330，配置为采用如下方式对行动器和评价器的参数进行更新：

设置折扣因子等价式；

将折扣因子等价式代入预设损失函数中计算，采用预设损失函数更新评判器在线网络 Q 的参数；

采用下式更新行动器的在线策略网络 μ ，

$$\nabla_{\theta^{\mu}} J = \frac{1}{m} \sum_n \nabla_a Q(s, a | \theta^Q) \big|_{s=s_n, a=\mu(s_n)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) \big|_{s_n}$$

其中，m 为随机抽取的样本数量， $\nabla_{\theta^{\mu}} \mu$ 表示对策略网络 u 的网络参数求偏导数， ∇_a 表示是对输入的动作 a 求偏导数； θ^{μ} 表示在线对策略网络 u 的参数； θ^Q 为评判器在线网络 Q 的参数；

5 采用软更新方式，更新行动器目标网络 μ' 和评判器目标网络 Q' 。

在本申请的一些可选实施例中，所述预设损失函数采用如下公式：

$$L = \frac{1}{m} \sum_n (y_n - Q(s_n, a_n | \theta^Q))^2$$

其中， y_n 表示折扣因子等价式； $y_n = r_n + \gamma Q'(s_{n+1}, \mu'(s_{n+1} | \theta^{\mu'}) | \theta^Q)$ ， $n = 1, 2 \dots m$ ；

θ^Q 为评判器在线网络 Q 的参数； S_n 为当前状态变量， a_n 为调度策略。

10 在本申请的一些可选实施例中，所述状态变量，表示各时刻系统的能流分布；

所述动作变量，表示系统的控制变量，包括常规机组发电功率、燃气轮机功率、电转气功率、储能充放电功率中的至少之一；

所述回报函数，表示每时刻系统的优化目标分类；

15 所述记忆库容量，表示强化学习智能体的记忆库容量，当智能体存储的 5 元组个数超过该容量值时，则随机删除记忆库中的某一个 5 元组，存储新的 5 元组。

在本申请的一些可选实施例中，所述装置还包括调度策略生成模块，配置为获取未来调度周期内的新能源发电功率以及电负荷、天然气负荷需求，在所述未来调度周期内的每一时间步，向所述调度模型的行动器中输入系统的运行状态，生成各时刻的调度策略。

需要说明的是：上述实施例提供的综合能源系统调度模型构建装置在进行模型构建时，仅以上述各程序模块的划分进行举例说明，实际应用中，可以根据需要而将上述处理分配由不同的程序模块完成，即将装置的内部

结构划分成不同的程序模块，以完成以上描述的全部或者部分处理。另外，上述实施例提供的综合能源系统调度模型构建装置与综合能源系统调度模型构建方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

5 本申请实施例还提供一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行本申请实施例综合能源系统调度模型构建方法。

其中，存储介质可以是任何的各种类型的存储器电子设备或存储电子设备。术语“存储介质”旨在包括：安装介质，例如只读光盘（CD-ROM，
10 Compact Disc Read-Only Memory）、软盘或磁带装置；计算机系统存储器或随机存取存储器，诸如动态随机存取存储器（DRAM，Dynamic Random Access Memory）、双倍数据速率同步动态随机存取存储器（DDRSDRAM，Double Data Rate Synchronous Dynamic Random Access Memory）、静态随机存取存储器（SRAM，Static Random Access Memory）等；非易失性存储器，
15 诸如闪存、磁介质（例如硬盘或光存储）；寄存器或其它相似类型的存储器元件等。存储介质可以还包括其它类型的存储器或其组合。另外，存储介质可以位于程序在其中被执行的计算机系统中，或者可以位于不同的第二计算机系统中，第二计算机系统通过网络（诸如因特网）连接到计算机系统。第二计算机系统可以提供程序指令给计算机用于执行。术语“存储介质”可
20 以包括可以驻留在不同位置中（例如在通过网络连接的不同计算机系统中）的两个或更多存储介质。存储介质可以存储可由一个或多个处理器执行的程序指令（例如具体实现为计算机程序）。

当然，本申请实施例所提供的一种包含计算机可执行指令的存储介质，其计算机可执行指令不限于如上所述的综合能源系统的调度操作，还可以
25 执行本申请任意实施例所提供的综合能源系统的调度方法中的相关操作。

本申请实施例提供了一种电子设备，该电子设备中可集成本申请实施例提供的综合能源系统调度模型构建装置。图 4 是本申请实施例提供的一种电子设备的结构示意图。如图 4 所示，本实施例提供了一种电子设备 400，其包括：一个或多个处理器 420；存储装置 410，用于存储一个或多个程序，
5 当所述一个或多个程序被所述一个或多个处理器 420 执行，使得所述一个或多个处理器 420 实现本申请任意实施例所提供的综合能源系统调度模型构建方法。

图 4 显示的电子设备 400 仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。

10 如图 4 所示，该电子设备 400 包括处理器 420、存储装置 410、输入装置 430 和输出装置 440；电子设备中处理器 420 的数量可以是一个或多个，图 4 中以一个处理器 420 为例；电子设备中的处理器 420、存储装置 410、输入装置 430 和输出装置 440 可以通过总线或其他方式连接，图 4 中以通过总线 450 连接为例。

15 存储装置 410 作为一种计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块单元，如本申请实施例中的综合能源系统调度模型构建方法对应的程序指令。

存储装置 410 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序；存储数据区可存储根
20 据终端的使用所创建的数据等。此外，存储装置 410 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实例中，存储装置 410 可进一步包括相对于处理器 420 远程设置的存储器，这些远程存储器可以通过网络连接。上述网络的实例包括但不限于互联网、企业内部网、局域网、
25 移动通信网及其组合。

输入装置 430 可用于接收输入的数字、字符信息或语音信息，以及产生与电子设备的用户设置以及功能控制有关的键信号输入。输出装置 440 可包括显示屏、扬声器等电子设备。

上述本发明实施例揭示的方法可以应用于处理器中，或者由处理器实现。处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器、数字信号处理器(DSP, Digital Signal Processor)，或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。处理器可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。结合本发明实施例所公开的方法的步骤，可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于存储介质中，该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成前述方法的步骤。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得

通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

5 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

10 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

15 由技术常识可知，本申请可以通过其它的不脱离其精神实质或必要特征的实施方案来实现。因此，上述公开的实施方案，就各方面而言，都只是举例说明，并不是仅有的。所有在本申请范围内或在等同于本申请的范围内的改变均被本申请包含。

权利要求书

1、一种综合能源系统调度模型构建方法，包括：

获取新能源发电功率以及电负荷、天然气负荷的历史数据，作为训练样本，并获取综合能源系统的优化目标；

5 基于所述训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素；

基于所述训练元素对深度强化学习模型进行迭代训练，生成调度模型。

2、根据权利要求 1 所述的方法，其中，所述优化目标包括：运行成本最小、碳排放量最小、新能源消纳率最高中的至少一种。

10 3、根据权利要求 1 所述的方法，其中，所述深度强化学习模型包括：行动器和评判器；所述基于所述训练元素对所述深度强化学习模型进行迭代训练，包括：

对构建的深度强化学习模型进行初始化；所述初始化包括设置训练迭代次数和每一次迭代的调度周期；

15 从调度周期内的第一个时刻开始，计算该时刻的综合能源系统的能流分布，得到系统环境的状态变量，并利用回报函数计算此刻系统的运行成本；

将所述状态变量输入至深度强化学习模型的行动器中，输出调度策略作为动作变量；

20 所述调度策略叠加一个正态分布的噪声后，作用于综合能源系统，得到下一状态变量；

将所述状态变量、所述动作变量、运行成本、所述下一状态变量以及是否为终止状态的 5 元组，存入强化学习智能体的经验回放单元；并从强化学习智能体的经验回放单元中抽取预设数量的样本，基于抽取的样本对
25 行动器和评价器的参数进行更新，持续迭代，直至迭代收敛稳定，得到训

练好的深度强化学习模型。

4、根据权利要求3所述的方法，其中，所述行动器包括行动器在线策略网络及其目标策略网络；所述评判器包括评判器在线网络及其目标网络；所述基于抽取的样本对行动器和评价器的参数进行更新，包括：

5 设置折扣因子等价式；

将折扣因子等价式代入预设损失函数中计算，采用预设损失函数更新评判器在线网络Q的参数；

采用下式更新行动器的在线策略网络 μ ：

$$\nabla_{\theta^{\mu}} J = \frac{1}{m} \sum_n \nabla_a Q(s, a | \theta^Q) \big|_{s=s_n, a=\mu(s_n)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) \big|_{s_n}$$

10 其中，m为随机抽取的样本数量， $\nabla_{\theta^{\mu}} \mu$ 表示对策略网络u的网络参数求偏导数， ∇_a 表示是对输入的动作a求偏导数； θ^{μ} 表示在线对策略网络u的参数； θ^Q 为评判器在线网络Q的参数；

采用软更新方式，更新行动器目标网络 μ' 和评判器目标网络Q'。

15 5、根据权利要求4所述的方法，其中，所述预设损失函数采用如下公式：

$$L = \frac{1}{m} \sum_n (y_n - Q(s_n, a_n | \theta^Q))^2$$

其中， y_n 表示折扣因子等价式； $y_n = r_n + \gamma Q'(s_{n+1}, \mu'(s_{n+1} | \theta^{\mu'}) | \theta^{Q'})$ ， $n=1, 2, \dots, m$ ；

θ^Q 为评判器在线网络Q的参数； S_n 为当前状态变量， a_n 为调度策略。

6、根据权利要求1所述的方法，其中，所述训练元素中：

20 所述状态变量，表示各时刻系统的能流分布；

所述动作变量，表示系统的控制变量，包括常规机组发电功率、燃气轮机功率、电转气功率、储能充放电功率中的至少之一；

所述回报函数，表示每时刻系统的优化目标分类；

所述记忆库容量，表示强化学习智能体的记忆库容量，当智能体存储

的 5 元组个数超过该容量值时，则随机删除记忆库中的某一个 5 元组，存储新的 5 元组。

7、根据权利要求 1 所述的方法，其中，所述方法还包括：

5 获取未来调度周期内的新能源发电功率以及电负荷、天然气负荷需求，
在所述未来调度周期内的每一时间步，向所述调度模型的行动器中输入系统的运行状态，生成各时刻的调度策略。

8、一种综合能源系统调度模型构建装置，包括：

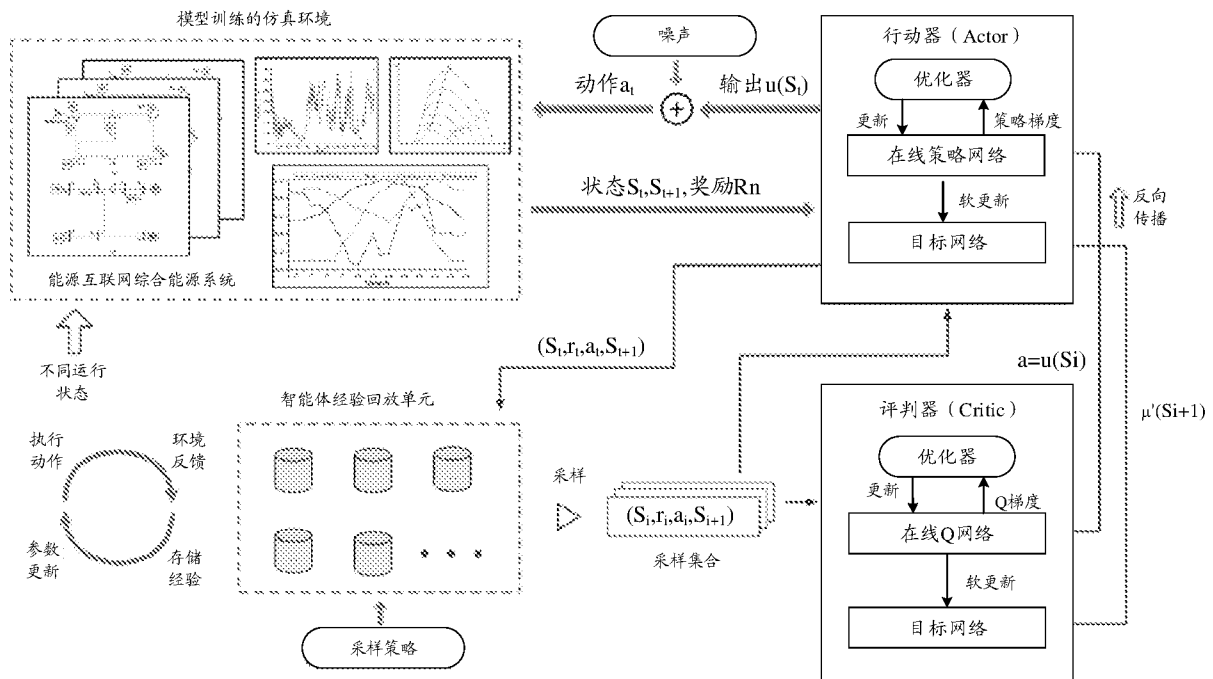
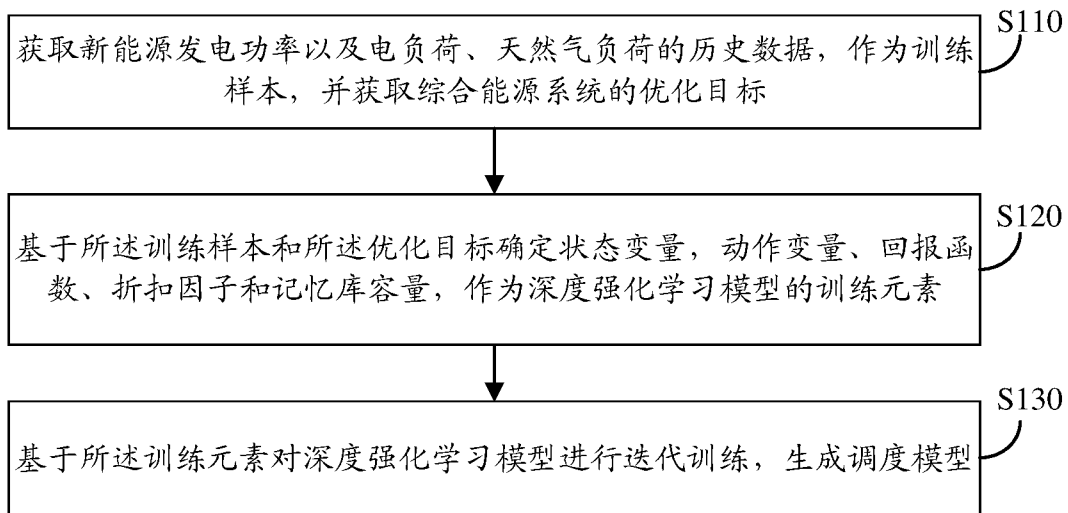
训练样本收集模块，配置为获取新能源发电功率以及电负荷、天然气负荷的历史数据，作为训练样本，并获取综合能源系统的优化目标；

10 模型构建模块，配置为基于所述训练样本和所述优化目标确定状态变量、动作变量、回报函数、折扣因子和记忆库容量，作为深度强化学习模型的训练元素；

调度策略迭代模块，配置为基于所述训练元素对深度强化学习模型进行迭代训练，生成调度模型。

15 9、一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如权利要求 1-7 中任一项所述的综合能源系统调度模型构建方法。

10、一种电子设备，包括存储器，处理器及存储在存储器上并可在处理器运行的计算机程序，所述处理器执行所述计算机程序时实现如权利要求 1-7 中任一项所述的综合能源系统调度模型构建方法。
20



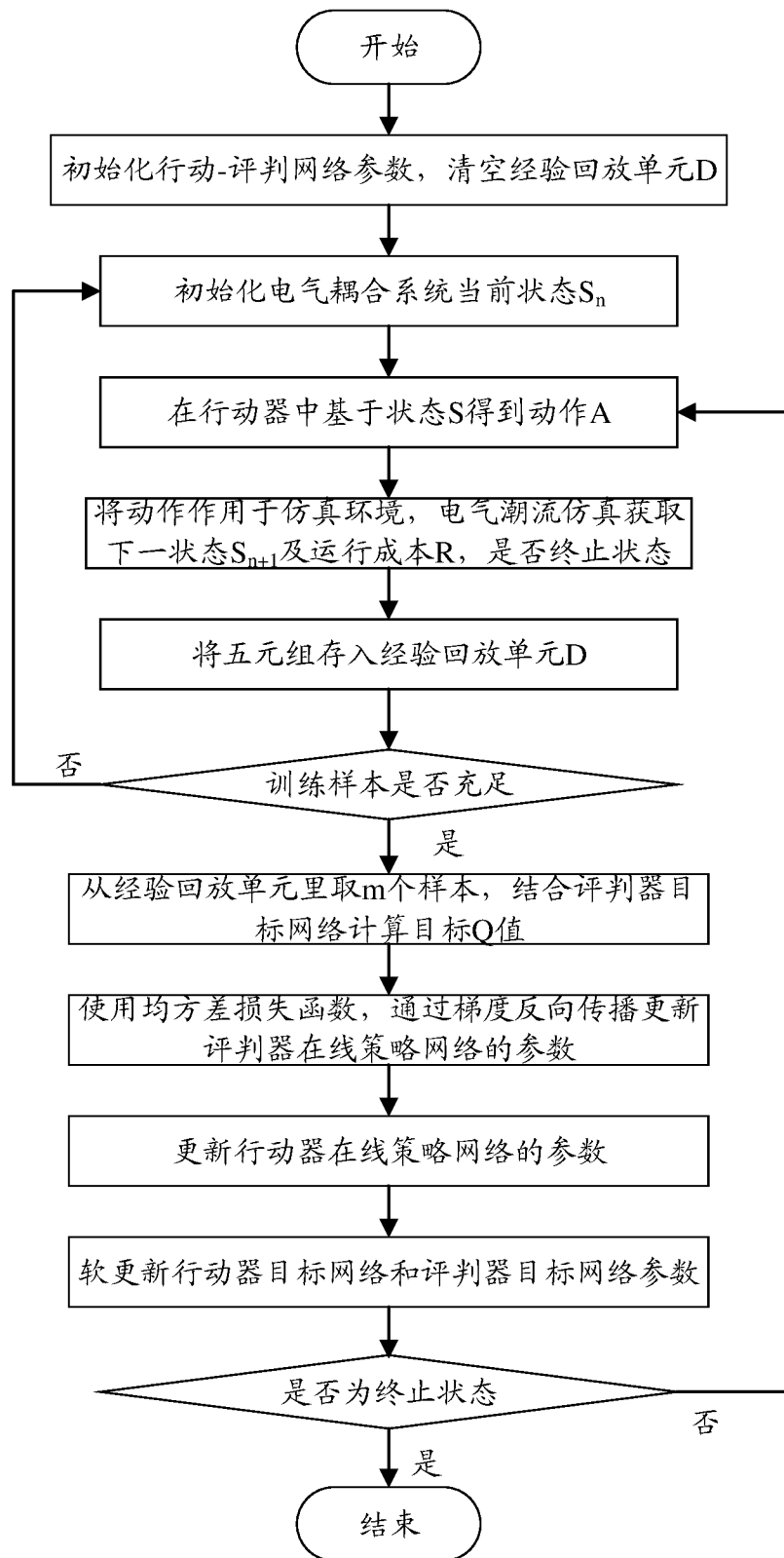


图 2 (b)

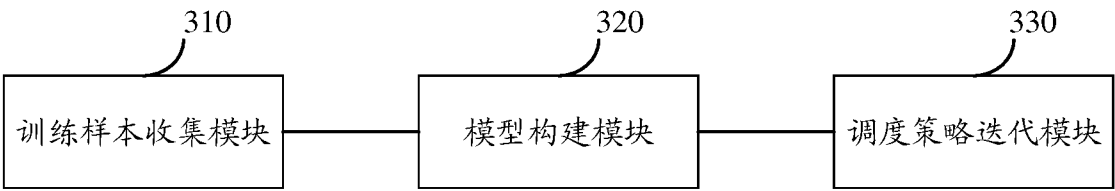


图 3

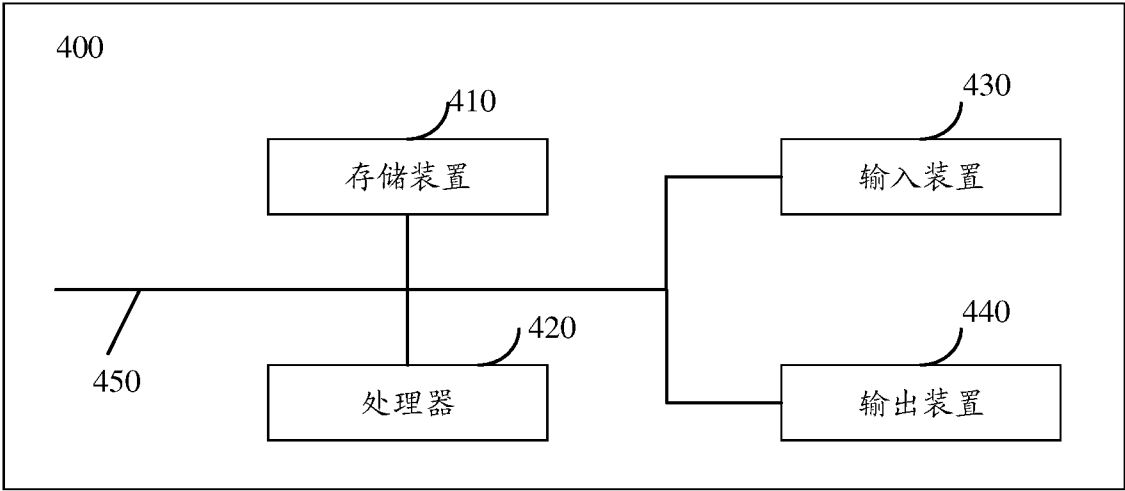


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/115751

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/06(2012.01)i; G06Q 50/06(2012.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI: 综合能源, 调度, 优化, 强化学习, 成本, 最小, 评价, 行动, integrated energy system, reinforcement learning, dispatch, optimi+, deep deterministic policy gradient, actor, critic, DDPG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	杨挺 等 (YANG, Ting et al.). "基于深度强化学习的综合能源系统动态经济调度 (Dynamic Economic Dispatch for Integrated Energy System Based on Deep Reinforcement Learning)" 电力系统自动化 (Automation of Electric Power Systems), 12 November 2020 (2020-11-12), ISSN: 1000-1026, pp. 1-10	1-10
A	CN 111064229 A (GUANGDONG UNIVERSITY OF TECHNOLOGY) 24 April 2020 (2020-04-24) entire document	1-10
A	CN 112214883 A (CHINA NATIONAL OFFSHORE OIL CORPORATION et al.) 12 January 2021 (2021-01-12) entire document	1-10
A	CN 111738529 A (ELECTRIC POWER RESEARCH INSTITUTE, CHINA SOUTHERN POWER GRID et al.) 02 October 2020 (2020-10-02) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 November 2021

Date of mailing of the international search report

25 November 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/115751

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	乔骥 等 (QIAO, Ji et al.). "基于柔性行动器-评判器深度强化学习的电-气综合能源系统优化调度 (Optimal Dispatch of Integrated Electricity-gas System With Soft Actor-critic Deep Reinforcement Learning)" <i>中国电机工程学报 (Proceedings of the CSEE)</i> , 17 November 2020 (2020-11-17), ISSN: 0258-8013, pp. 1-15	1-10
A	WANG, Yiting et al. "Energy Management of Integrated Energy System Based on Stackelberg Game and Deep Reinforcement Learning" <i>Preprints of the 4th IEEE Conference on Energy Internet and Energy System Integration</i> , 01 November 2020 (2020-11-01), pp. 2645-2651	1-10
PX	CN 112862281 A (CHINA ELECTRIC POWER RESEARCH INSTITUTE CO., LTD. et al.) 28 May 2021 (2021-05-28) claims 1-10	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/115751

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111064229	A	24 April 2020	None	
CN	112214883	A	12 January 2021	None	
CN	111738529	A	02 October 2020	None	
CN	112862281	A	28 May 2021	None	

国际检索报告

国际申请号

PCT/CN2021/115751

A. 主题的分类 G06Q 10/06(2012.01)i; G06Q 50/06(2012.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06Q 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) WPI, EPDOC, CNPAT, CNKI:综合能源, 调度, 优化, 强化学习, 成本, 最小, 评价, 行动, integrated energy system, reinforcement learning, dispatch, optimi+, deep deterministic policy gradient, actor, critic, DDPG		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	杨挺 等. "基于深度强化学习的综合能源系统动态经济调度" 电力系统自动化, 2020年 11月 12日 (2020 - 11 - 12), ISSN: 1000-1026, 第1-10页	1-10
A	CN 111064229 A (广东工业大学) 2020年 4月 24日 (2020 - 04 - 24) 全文	1-10
A	CN 112214883 A (中国海洋石油集团有限公司 等) 2021年 1月 12日 (2021 - 01 - 12) 全文	1-10
A	CN 111738529 A (南方电网科学研究院有限责任公司 等) 2020年 10月 2日 (2020 - 10 - 02) 全文	1-10
A	乔骥 等. "基于柔性行动器-评判器深度强化学习的电-气综合能源系统优化调度" 中国电机工程学报, 2020年 11月 17日 (2020 - 11 - 17), ISSN: 0258-8013, 第1-15页	1-10
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件		
国际检索实际完成的日期 2021年 11月 15日		国际检索报告邮寄日期 2021年 11月 25日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国 北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 方蕾 电话号码 86-(10)-53961493

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	WANG, Yiting 等. "Energy Management of Integrated Energy System Based on Stackelberg Game and Deep Reinforcement Learning" Preprints of the 4th IEEE Conference on Energy Internet and Energy System Integration, 2020年 11月 1日 (2020 - 11 - 01), 第2645-2651页	1-10
PX	CN 112862281 A (中国电力科学研究院有限公司 等) 2021年 5月 28日 (2021 - 05 - 28) 权利要求1-10	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/115751

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	111064229	A	2020年 4月 24日	无	
CN	112214883	A	2021年 1月 12日	无	
CN	111738529	A	2020年 10月 2日	无	
CN	112862281	A	2021年 5月 28日	无	