

(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
2024 年 7 月 25 日 (25.07.2024)



(10) 国际公布号
WO 2024/152902 A1

(51) 国际专利分类号:
H04L 41/0894 (2022.01) **G06F 16/36** (2019.01)

(21) 国际申请号: PCT/CN2024/070341

(22) 国际申请日: 2024 年 1 月 3 日 (03.01.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310089923.8 2023年1月17日 (17.01.2023) CN

(71) 申请人: 烽火通信科技股份有限公司 (FIBERHOME TELECOMMUNICATION TECHNOLOGIES CO., LTD.) [CN/CN]; 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。

(72) 发明人: 徐安然(XU, Anran); 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。李文超(LI, Wenchao); 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。尹山(YIN, Shan); 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。匡立伟(KUANG, Liwei); 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。谢秋红(XIE, QiuHong); 中国湖北省武汉市东湖高新技术开发区高新四路6号, Hubei 430000 (CN)。

(74) 代理人: 深圳市六加知识产权代理有限公司(LIUJIA CHINA IP LAW OFFICE); 中国湖北省武汉市洪山区光谷大道3号未来之光5栋401室, Hubei 430000 (CN)。

(54) Title: REINFORCEMENT-LEARNING-BASED DIGITAL TWIN NETWORK SLICING METHOD AND APPARATUS

(54) 发明名称: 一种基于强化学习的数字孪生网络切片的方法和装置

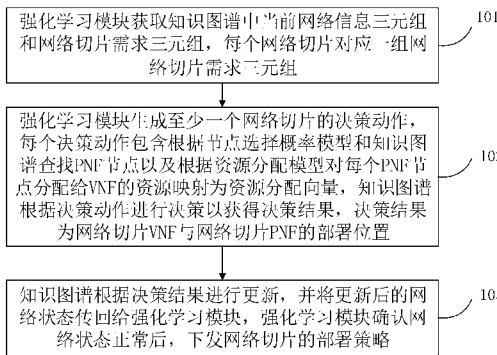


图 2

- 101 A reinforcement learning module acquires the current network information triple and a network slicing requirement triple from a knowledge graph, wherein each network slice corresponds to a network slicing requirement triple
- 102 The reinforcement learning module generates at least one decision-making action for network slicing, wherein each decision-making action involves: searching for a PNF node according to a node selection probability model and the knowledge graph, and mapping a resource, which is allocated to a VNF by each PNF node according to a resource allocation model, into a resource allocation vector; and the knowledge graph making a decision according to the decision-making action to obtain a decision-making result, wherein the decision-making result indicates deployment positions of a network slice VNF and a network slice PNF
- 103 The knowledge graph performs update according to the decision-making result, and transmits an updated network state back to the reinforcement learning module, and the reinforcement learning module confirms that the network state is normal and then issues a network slice deployment policy

(57) Abstract: The present invention relates to the field of network communications, and particularly relates to a reinforcement-learning-based digital twin network slicing method and apparatus. The method mainly comprises: a reinforcement learning module acquiring the current network information triple and a network slicing requirement triple from a knowledge graph; the reinforcement learning module generating at least one decision-making action for network slicing, wherein each decision-making action includes: searching for a PNF node according to a node selection probability model and the knowledge graph, and mapping a resource, which is allocated

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

to a VNF by each PNF node according to a resource allocation model, into a resource allocation vector; the knowledge graph making a decision according to the decision-making action to obtain a decision-making result, wherein the decision-making result indicates deployment positions of a network slice VNF and a network slice PNF; and the knowledge graph performing update according to the decision-making result, and transmitting an updated network state back to the reinforcement learning module, and the reinforcement learning module issuing a network slice deployment policy. By means of the present invention, the optimal slice deployment scheme in the current network scenario can be quickly obtained in a dynamic network environment.

(57) 摘要: 本发明涉及网络通信领域, 特别是涉及一种基于强化学习的数字孪生网络切片的方法和装置。主要包括: 强化学习模块获取知识图谱中当前网络信息三元组和网络切片需求三元组; 强化学习模块生成至少一个网络切片的决策动作, 每个决策动作包含根据节点选择概率模型和知识图谱查找PNF节点以及根据资源分配模型对每个PNF节点分配给VNF的资源映射为资源分配向量, 知识图谱根据决策动作进行决策以获得决策结果, 决策结果为网络切片VNF与网络切片PNF的部署位置; 知识图谱根据决策结果进行更新, 并将更新后的网络状态传回给强化学习模块, 强化学习模块下发网络切片的部署策略。本发明可以在动态网络环境下快速获得当前网络场景中最优的切片部署方案。

一种基于强化学习的数字孪生网络切片的方法和装置

相关申请的交叉参考

本申请要求于如下六件专利申请的优先权：

- 5 (1)、2023 年 1 月 17 日提交中国专利局，申请号为 202310089923.8，名称为“一种基于强化学习的数字孪生网络切片的方法和装置”的中国专利申请。

技术领域

本发明涉及网络通信领域，特别是涉及一种基于强化学习的数字孪生网络切片的方法和装置。

背景技术

- 10 随着 5G 和云技术的发展，越来越多的新种类业务逐渐涌现出来，而不同的业务或用户对网络的服务质量要求不尽相同。为了满足用户们的多样化需求，提供个性化差异服务，网络切片技术应运而生。如图 1 所示，网络切片技术可以在一张共享的物理网络上灵活调配资源，按需构建多个专用的、互相隔离的虚拟化逻辑网络，即多个网络切片，每个网络切片用以满足特定的业务需求。
- 15 网络切片是通过网络虚拟化技术实现的，由一组满足业务需求的虚拟网络功能（Virtual Network Function, 简称为 VNF）以及连接每两个虚拟网络功能的链路组成。对网络进行切片可以看作是基于共享的网络基础设施进行 VNF 以及链路的资源分配过程。如何对有限的资源进行灵活高效的分配，以实现各个网络切片的服务需求是网络切片的热点研究问题。

- 20 当前网络切片的部署大多仅考虑节点资源能力大小，没有综合考虑资源能力、链路长度以及物理网络功能（physical network function, 简称为 PNF）运行状态，难以保证不同网络需求和网络资源的高效利用。并且，物理网络资源是在时刻变化的，并且同时会产生多个网络切片需求，当前的网络切片的部署和编排方法大多没有考虑到物理网络状态的动态变化以及多网络切片需求下切片部署策略的调整。
- 25

鉴于此，如何克服现有技术所存在的缺陷，解决网络切片时策略调整困难的现象，是本技术领域待解决的问题。

申请内容

- 30 针对现有技术的以上缺陷或改进需求，本发明解决了网络切片时策略调整困难的问题。

本发明实施例采用如下技术方案：

第一方面，本发明提供了一种基于强化学习的数字孪生网络切片的方法，具体为：强化学习模块获取知识图谱中当前网络信息三元组和网络切片需求三元组，每个网络切片对应一组网络切片需求三元组，其中，当前网络信息三元组包括网络中所有 PNF 连接关系和每个 PNF 的资源属性，网络切片需求三元组包括切片需求的所有 VNF 连接关系、每个 VNF 的资源需求和切片优先级；强化学习模块生成至少一个网络切片的决策动作，每个决策动作包含根据节点选择概率模型和知识图谱查找 PNF 节点以及根据资源分配模型对每个 PNF 节点分配给 VNF 的资源映射为资源分配向量，知识图谱根据决策动作进行决策以获得决策结果，决策结果为网络切片 VNF 与网络切片 PNF 的部署位置；知识图谱根据决策结果进行更新，并将更新后的网络状态传回给强化学习模块，强化学习模块确认网络状态正常后，下发网络切片的部署策略。

优选的，所述强化学习模块生成至少一个网络切片的决策动作，具体包括：根据切片需求，基于知识图谱，找到节点选择概率模型中概率最大的 PNF 作为第一个节点；基于知识图谱得到资源评估向量，输入到资源分配模型中为第一 VNF 节点分配资源，并更新知识图谱；根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到可能的下一个 VNF 节点的 PNF 位置选择，依次查找直至所有节点都部署分配完毕。

优选的，所述找到节点选择概率模型中概率最大的 PNF 作为第一个节点，具体包括：找到第一个 VNF 节点的动作空间中所有可能的 PNF 位置选择作为第一 VNF 节点的部署动作空间，并得到节点状态矩阵；将节点状态矩阵输入到节点选择概率模型中，计算每个节点被选择的概率，选择其中概率最大的 PNF 作为第一个节点；基于知识图谱得到第一个节点的资源评估向量，将评估向量输入到资源分配模型中，为第一 VNF 节点分配资源。

优选的，所述找到可能的下一个 VNF 节点的 PNF 位置选择，具体包括：根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到所有可能的下一个 VNF 节点的 PNF 位置选择，作为下一个 VNF 节点的部署动作空间；基于知识图谱得到节点状态矩阵，依次向后进行下一个节点的部署和资源分配，直至所有节点都部署分配完毕，以所有节点的分配结果完成一整个切片部署策略的选择。

优选的，所述知识图谱根据决策结果进行更新，具体包括：知识图谱根据强化学习的资源分配策略进行更新，并根据反馈函数计算反馈值；结合反馈值利用梯度下降法更新节点选择概率模型的参数。

优选的，所述节点选择概率模型具体包括：将 PNF 之间的距离、切片对于 VNF 节点的资源需求和 PNF 上的资源状态映射成可能选择的概率，以便于选择强化学习模块获取概率最大的 PNF 节点进行部署。

5 优选的，所述资源分配模型具体包括：通过函数将资源评估向量映射成资源分配向量，以便于表示 PNF 节点分配给 VNF 的计算、存储、网络资源。

10 优选的，所述强化学习模块确认网络状态正常，还包括：当网络切片请求在进行某个节点部署，由于之前的网络切片占用过多资源而无法找到资源满足需求的 PNF 时，根据网络切片的优先级排序，在不低于最低需求的条件下，按比例缩小前几个网络切片中每个 VNF 的资源，缩小的资源大小与网络切片的优先级成反比。

优选的，所述强化学习模块确认网络状态正常，还包括：当网络状态发生改变，知识图谱模块中当前时刻的网络状态无法满足原时刻的切片需求时，则强化学习模块退回到不满足的切片步骤重新进行部署决策。

15 另一方面，本发明提供了一种基于强化学习的数字孪生网络切片的装置，具体为：包括至少一个处理器和存储器，至少一个处理器和存储器之间通过数据总线连接，存储器存储能被至少一个处理器执行的指令，指令在被处理器执行后，用于完成第一方面中的基于强化学习的数字孪生网络切片的方法。

20 与现有技术相比，本发明实施例的有益效果在于：使用强化学习模块生成网络切片部署策略，在动态网络环境下快速完成时间切片部署方案的决策，获得当前网络场景中最优的切片部署方案，更好地保证网络资源的高效利用。

附图说明

25 为了更清楚地说明本发明实施例的技术方案，下面将对本发明实施例中所需要使用的附图作简单地介绍。显而易见地，下面所描述的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为执行网络切片的网络结构模型示意图；

图 2 为本发明实施例提供的一种基于强化学习的数字孪生网络切片的方法流程图；

30 图 3 为本发明实施例提供的另一种基于强化学习的数字孪生网络切片的方法流程图；

图 4 为本发明实施例提供的另一种基于强化学习的数字孪生网络切片的方法流程图；

图 5 为本发明实施例提供的强化学习模块架构示意图；

图 6 为本发明实施例提供的一种基于强化学习的数字孪生网络切片的装置结构示意图。

具体实施方式

为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本发明，并不用于限定本发明。

本发明是一种特定功能系统的体系结构，因此在具体实施例中主要说明各结构模组的功能逻辑关系，并不对具体软件和硬件实施方式做限定。

此外，下面所描述的本发明各个实施方式中所涉及到的技术特征只要彼此之间未构成冲突就可以相互组合。下面就参考附图和实施例结合来详细说明本发明。

实施例 1：

强化学习（Reinforcement Learning，简称为 RL）是一种人工智能技术，主要思想是基于和环境的持续交互，获得反馈，并据此不断优化决策。本专利通过结合知识图谱和强化学习来实现网络切片部署决策，可以更好地保证网络资源的高效利用。

针对以上问题，本专利所描述的方法通过一种基于强化学习的网络切片部署策略生成方法实现网络切片，生成的网络切片的部署策略可以更好地保证每个网络切片以及整个网络的运行状态。

如图 2 所示，本发明实施例提供的基于强化学习的数字孪生网络切片的方法具体步骤如下：

步骤 101：强化学习模块获取知识图谱中当前网络信息三元组和网络切片需求三元组，每个网络切片对应一组网络切片需求三元组。

强化学习模块的输入是来自知识图谱的，知识图谱由“实体-关系-实体”三元组以及实体及其相关属性-值对组成，三元组可表示为（Thing1，Relation，Thing2），Thing1 表示第一个实体，Thing2 表示第二个实体，Relation 表示 Thing1 和 Thing2 之间的关系。本实施例的方法用于根据当前网络状态计算网络切片的部署策略，因此需要使用网络信息三元组和网络切片需求三元组。当前网络信息三元组包括网络中所有 PNF 连接关系，和每个 PNF 的资源属性，资源属性包括每个 PNF 的计算、存储、网络等已确定的硬件资源属性或可使用的资源属性。网络信息三元组中，Thing1 表示 PNF，Relation 表示该三元组的

资源类型，Thing2 表示 Thing1 所需的该类型资源的值。例如在某个具体场景中，某个 PNF（表示为 PNF1）需要的计算、存储、网络资源分别为 100、200、300，PNF1 与 PNF2 之间存在连接关系，该 PNF 可表示为：(PNF1, 连接, PNF2)、(PNF1, 计算, 100)、(PNF1, 存储, 200)、(PNF1, 网络, 300)。网络切片需求三元组包括切片需求的所有 VNF 连接关系、每个 VNF 的资源需求和切片优先级。每个网络切片对应一组网络切片需求三元组。网络切片需求三元组中，Thing1 表示网络切片或 VNF，Relation 表示该三元组的资源类型，Thing2 表示 Thing1 所需的该类型资源的值。进一步的，当 Thing1 为网络切片时，Relation 的类型为“所需的 VNF”或优先级等，Thing2 为 Thing1 需要的 VNF 或优先级值，例如，网络切片 1 中需要 VNF1，三元组为（切片 1，需要，VNF1）。当 Thing1 为 VNF 时，与网络信息三元组类似，Relation 为 VNF 分配的计算、存储、网络等资源类型，Thing2 为该资源类型的值。例如，VNF1 的计算资源分配值为 100，VNF1 的计算资源三元组为（VNF1，计算，100）。在实际使用中，强化学习模块同时会收到多个网络切片的需求。

步骤 102：强化学习模块生成至少一个网络切片的决策动作，每个决策动作包含根据节点选择概率模型和知识图谱查找 PNF 节点以及根据资源分配模型对每个 PNF 节点分配给 VNF 的资源映射为资源分配向量，知识图谱根据决策动作进行决策以获得决策结果，决策结果为网络切片 VNF 与网络切片 PNF 的部署位置。

强化学习模块的输出是网络切片的部署策略，即一个网络切片中每个 VNF 部署在哪个 PNF 上，每个 VNF 被分配多少资源。强化学习模块的输出会传到知识图谱，知识图谱进行更新。

步骤 103：知识图谱根据决策结果进行更新，并将更新后的网络状态传回给强化学习模块，强化学习模块确认网络状态正常后，下发网络切片的部署策略。

步骤 102 中，通过强化学习模块的决策，网络切片需要的 VNF 会改变，每个 VNF 所需的计算、存储、网络等资源也会出现变化，因此，需要根据决策后的数据对知识图谱中的三元组进行相应更新。例如，网络切片 1 中原本只需要 VNF1，表示为（切片 1，需要，VNF1），决策后变化为需要 VNF1 和 VNF2，三元组更新为（切片 1，需要，VNF1）、（切片 1，需要，VNF2）。又例如，VNF1 的计算资源分配值由 100 变为 150，相应的三元组也需要由（VNF1，计算，100）更新为（VNF1，计算，150）。

本实施例提供的方法中，将一个网络切片需求转化为多个 VNF 的分配需求，每个 VNF 的分配过程，都是先确认分配在哪个 PNF 上，然后再根据需求分配相应的资源。每分配完成一个 VNF 都会出现一次三元组的变化，需要根据新的三元组对下一个 VNF 进行却侧，因此，每次分配完成一个 VNF 后，均
5 有一次与环境知识图谱的交互。

知识图谱将更新后的网络状态传回强化学习模块，强化学习模块对调整后的网络状态做出评估，确认使用该切片策略后网络状态正常后，将部署决策信息传至控制模块进行下发。

经过本实施例中提供的步骤 101-步骤 103 后，即可根据当前网络状态，使用强化学习模块完成网络切片的部署决策最优化选择，并对决策结果进行确认，
10 确保执行新的切片策略后网络状态正常。

基于强化学习进行网络切片部署决策主要由以下几部分组成：环境、动作、状态、奖励和决策模型。强化学习的环境为网络知识图谱。状态是指网络当前的状态。动作是强化学习在当前状态下做出的一个决策，即一个网络切片的节点选择方案和资源分配方案。奖励是指强化学习做出一个动作后影响环境，环境所给的反馈，反映出切片部署决策的好坏。决策模型在本专利中包括节点选择概率模型和资源分配模型，强化学习根据决策模型产生一个切片部署的决策动作。
15

如图 3 所示，可以使用以下步骤完成强化学习模块生成网络切片的部署策略。
20

步骤 201：根据切片需求，基于知识图谱，找到节点选择概率模型中概率最大的 PNF 作为第一个节点。

根据切片需求，基于知识图谱，找到第一个 VNF 节点的动作空间中所有可能的 PNF 位置选择作为第一 VNF 节点的部署动作空间，并得到节点状态矩阵。将节点状态矩阵输入到节点选择概率模型中，计算每个节点被选择的概率，选择其中概率最大的 PNF 作为第一个节点。再基于知识图谱得到第一个节点的资源评估向量，将评估向量输入到资源分配模型中，为第一 VNF 节点分配资源。
25

步骤 202：基于知识图谱得到资源评估向量，输入到资源分配模型中为第一 VNF 节点分配资源，并更新知识图谱。
30

完成一个节点的部署动作后，知识图谱根据强化学习的资源分配策略进行更新，并根据反馈函数计算反馈值。

在进行更新时，还需要结合反馈值利用梯度下降法更新节点选择概率模型的参数。具体的计算公式如下。

$$w = w + \eta * \frac{\partial L(w)}{\partial w} * r \quad (1)$$

该公式表示，节点选择概率模型中，每个节点的选择概率根据反馈值进行梯度下降后，当前的节点选择概率模型。

步骤 203：根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到可能的下一个 VNF 节点的 PNF 位置选择，依次查找直至所有节点都部署分配完毕。

完成第一个 VNF 节点选择后，还需要依次向后获取所有 VNF 节点的 PNF 位置。根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到所有可能的下一个 VNF 节点的 PNF 位置选择，作为下一个 VNF 节点的部署动作空间。再基于知识图谱得到节点状态矩阵，重复前面步骤，依次向后进行下一个节点的部署和资源分配，直至所有节点都部署分配完毕，以所有节点的分配结果完成一整个切片部署策略的选择。

在实际决策过程中，若在进行第 k 个节点的部署策略选择时，发现没有可供选择的节点位置选项，则退回到第 k-1 个节点，选择概率第二大的选项再进行下一步。

当需要进行多个网络切片时，对一个网络切片请求生成部署策略后，进行下一个网络切片的部署决策。

经过本实施例中提供的步骤 201-步骤 203 后，即可完成步骤 102 中网络切片的决策，获取到所有 VNF 和 PNF 的部署策略。与前述步骤相结合后，网络切片和调整的完整过程如图 4 所示。

进一步的，根据上述步骤完成所有网络切片部署决策后，还需要将部署决策传到知识图谱模块，对知识图谱模块进行更新。知识图谱再将更新后的网络状态传回强化学习模块。通过循环反馈的过程提高决策准确度，并使知识图谱中的对应关系能够在网络环境出现动态变化时与当前环境相符。

在实际实施场景中，可以通过以下方式构建强化学习模型，并完成部署和反馈。具体实施中，可以使用如图 5 所示的强化学习模块。

网络切片的网络运行状态具体包括网络切片 PNF 和网络切片 VNF 的位置、资源和关系信息。本实施例中，使用以下数学模型对网络运行状态进行描述：

$N^P = \{n_1^P, n_2^P, \dots\}$ 表示物理集群层中的 PNF。 $l^P = \{l_{1,2}^P, l_{1,3}^P, \dots\}$ 表示物理集群层中的 PNF 与 PNF 之间的物理链路，例如， $l_{ij}^P = (n_i^P, n_j^P)$ 表示切片中 PNF n_i^P 与 PNF n_j^P 之间的物理链路。网络中存在的网络切片用 $S = \{S_1, S_2, \dots, S_b\}$ 表示，总计有 b 个网络切片。在一个网络切片 S_s 中， $N^S = \{n_1^S, n_2^S, \dots\}$ 表示 VNF 节点，

5 $l_s = \{l_{1,2}^S, l_{1,3}^S, \dots\}$ 表示 VNF 节点之间的虚拟链路，例如 $l_{ij}^S = (n_i^S, n_j^S)$ 表示切片中 VNF n_i^S 与 VNF n_j^S 之间的虚拟链路。对于 PNF n_i^P 的计算、存储、网络资源分别表示为 $H_{n_i^P}^1, H_{n_i^P}^2, H_{n_i^P}^3$ ，切片中 VNF 节点 n_i^S 的计算、存储、网络资源分别表示为 $h_{n_i^S}^1, h_{n_i^S}^2, h_{n_i^S}^3$ 。

每个 PNF 或 VNF 可靠性由失效平均时间 (Mean Time To Failure, 简称为

10 MTTF) 衡量，可服务性由平均修复时间 (Mean Time To Restoration, 简称为 MTTR) 衡量，则 PNF (VNF) 的可用性 A_{PNF} (A_{VNF}) 可以使用以下公式表示。

$$A_{PNF} = \frac{MTTF}{MTTF + MTTR} \quad (2)$$

整个切片的可用性等于构成切片 PNF 和 VNF 的可用性最小值，用于评估网络运行状态。

15 VNF 和 PNF 的位置关系用 $a_{n_i^S n_i^P}$ 可以用以下公式表示：

$$a_{n_i^S n_i^P} = \begin{cases} 1 & \text{if } n_i^S \text{ 在 } n_i^P \text{ 上} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

上述公式表示：如果 VNF n_i^S 在 PNF n_i^P 上，则 $a_{n_i^S n_i^P} = 1$ ，否则为 0。

VNF n_i^S 和 PNF n_i^P 的计算、存储和网络资源应满足以下条件：

$$\forall q=1,2,3 \sum_{s=1}^b h_{n_i^S}^q a_{n_i^S n_i^P} < H_{n_i^P}^q \quad (4)$$

20 上述公式表示：根据 PNF、VNF 可用性以及资源条件定义网络运行态，具体如下：对网络切片时，选择 VNF 并选择所部署的 PNF，按照空间位置关系对 PNF 排序。

本实施例的一个具体实施场景中，PNF 的总数为 k 个，所有 PNF 的集合表示为 $PNF(PNF_1, PNF_2, \dots, PNF_k)$ ；VNF 的总数为 r 个，所有 VNF 的集合表示为

25 $VNF(VNF_1, VNF_2, \dots, VNF_r)$ 。

PNF/VNF 初始状态由四部分组成：第一部分表示前一个 PNF/VNF 序号，

第一个 PNF/VNF 首位为 0。第二部分表示 PNF/VNF 所在 PNF 序号，第一个 PNF/VNF 第二部分表示为 1。前两部分的比特数会根据 PNF/VNF 的数量大小变化，例如总共有 4 个 PNF，则用 2 个比特表示，若有 10 个 PNF，比特数则需要为 4。第三部分表示 PNF/VNF 中计算、存储、网络资源。对于 PNF，1 表示计算/存储/网络资源有剩余，0 表示已用完；对于 VNF，1 表示在当前 PNF 上计算/存储/网络资源有剩余，0 表示已用完。第四部分表示，PNF/VNF 的运行状态，通过上面定义的可用性来判断，当计算得到的可用性高于设定的可用性阈值时，则是正常运行为 1，非正常运行为 0。例如，当有 2 个 PNF、3 个 VNF 时，其中 VNF1，VNF2 部署在 PNF1 上，VNF3 部署在 PNF2 上，均运行正常，资源充足。对于每个状态值，前两部分最大值为 3（3 个 VNF）分别占用 2 个 bit，第三部分占用 3 个 bit，第四部分占用 1 个 bit。则 PNF 初始状态表示分别为 00011111,01101111。VNF 初始状态分别为 00011111,01011111,10101111。例如，PNF 初始状态 00011111，自首位开始：第一部分 00 表示前一个 PNF 的序号；第二部分 01 表示所在 PNF 的序号；第三部分 111 中的每一位分别对应 VNF 的计算、存储、网络资源，该位为 1 表示对应资源有剩余；第四部分 1 表示 VNF 的运行状态正常。

强化学习模块中，网络切片部署的动作空间即切片中一个节点可以选择的节点部署策略集合。一个动作表示一个节点部署策略，可以表示为 $A = (a_1, a_2)$ 。其中 a_1 表示选择一个 PNF 节点， a_2 表示该节点分配的资源数量。部署一个节点的可选的动作有 k 个，则动作空间可以表示为 $A^S = (A_1, A_2, \dots, A_k)$ 。动作空间会受到网络资源限制的约束。

节点选择概率模型用于选择一个节点部署动作。节点选择概率模型将 PNF 之间的距离、切片对于 VNF 节点的资源需求和 PNF 上的资源状态映射成可能选择的概率，以便于选择强化学习模块获取概率最大的 PNF 节点进行部署。

节点选择概率模型的输入是一个节点状态矩阵 X 。矩阵的每一行为每个 PNF 节点的节点状态向量 $X_i = (X_i^1, X_i^2, X_i^3, X_i^4, X_i^5)$ 。其中， X_i^1 表示 PNF 节点 i 的可选性，根据前几跳是否被选择过以及是否可达来判断，可选为 1，不可选为 0。 X_i^2 表示 PNF 节点 i 到上一个 PNF 节点的可用的最短链路的物理距离，若为第一个节点，则 X_i^2 都为 1，若 PNF 节点 i 到上一个 PNF 节点没有链路可达，则 X_i^2 都为 -1。 X_i^3, X_i^4, X_i^5 分别表示 PNF 节点 i 上计算、存储和网络资源对需求的满足程度，通过计算 PNF 节点 i 上的可用资源与切片对该节点的最小资源需求的

差值得到。假设一个节点的资源需求是 (10,10,10)， PNF_i 的可用资源为 (100,100,90)，前几跳没有被选用过， PNF_i 到上一个 PNF 节点的物理距离为 3，则 PNF_i 对应的节点状态向量表示为 $X_i = [1, 3, 90, 90, 80]$ 。

节点概率选择模型先对节点状态矩阵进行加权计算： $Y = WX + b$ 得到每一个 PNF 节点是否可以作为下一个 PNF 节点的评价分数 $Y = (y_1, y_2, \dots, y_n)$ ，然后根据评价分数计算每个 PNF 节点作为被选择的概率 $P_i = \frac{y_i}{\sum_{k=0}^N y_k}$ 。本专利使用神经网络来参数化节点选择概率模型，其中 W 为神经元权重参数， b 为模型偏置参数，均由事先的模型训练得到。

资源分配模型的输入是一个资源评估向量 $x_r = (x_r^1, x_r^2, x_r^3, x_r^4, x_r^5, x_r^6, x_r^7)$ 。其中 x_r^1, x_r^2, x_r^3 表示 VNF 计算、存储、网络资源的最小需求， x_r^4, x_r^5, x_r^6 表示 PNF 节点上计算、存储、网络资源的可用量， x_r^7 表示网络切片的优先级。资源模型通过函数将资源评估向量映射成资源分配向量，资源分配向量可以表示为： $y_r = (y_r^1, y_r^2, y_r^3)$ ，其中， y_r^1, y_r^2, y_r^3 分别表示 PNF 节点分配给 VNF 的计算、存储、网络资源，在具体实施中，可以根据实际需要选择合适的参数作为资源分配向量的参数。资源分配模型表示为 $y_r = f(x_r)$ ，映射函数 $f(\cdot)$ 可以通过神经网络拟合。

反馈函数用于评判一个动作及部署策略的好坏，给出奖励。强化学习模型在选择策略动作时，会根据上次动作选择得到的奖励更新策略模型。选择奖励最高的切片部署策略就是强化学习的目标。反馈函数由三部分组成，包括部署精度、切片需求满足度和链路效率。部署精度表示所选择的 VNF、PNF 和链路是否能正常工作。例如，当 VNF 选取节点资源使用已超过所在 PNF 资源限制时，部署精度为 0，当选取的 VNF、PNF 运行不正常时，部署精度为 0。

$$r_{\text{ACCURACY}} = \begin{cases} -1, & \text{所选择的 VNF/PNF 不能正常工作} \\ +1, & \text{otherwise} \end{cases} \quad (5)$$

切片需求满足度指分配给切片的资源满足程度。考虑两种情况，一种情况是当分配给切片的计算、存储和网络资源不足以满足业务要求。另一种情况是分配给切片的资源可以满足要求，但可能分配的资源过多导致网络资源利用效率低。因此切片需求满足程度可以由以下公式表示。

$$r_{SATISFACTION} = \begin{cases} \alpha \cdot e^x, & \text{分配的资源不满足要求} \\ \gamma \frac{u_d}{a_d}, & \text{otherwise} \end{cases} \quad (6)$$

其中， α 表示不满足需求时的反馈参数， x 表示一个节点不满足的资源数量， γ 表示资源分配过多时的反馈参数， u_d 表示一个节点 VNF 的资源利用率， a_d 表示分配给这个节点的资源。

5 链路效率反应的是两个相邻的 VNF 所部署的 PNF 的距离远近。两个相邻 VNF 所在 PNF 节点应该尽可能相邻，若距离过远时延会增加。链路效率可以用以下公式表示。

$$r_{EFFICIENCY} = \frac{1}{\text{dist}(\text{VNF}_i, \text{VNF}_j)} \quad (7)$$

综合三部分后反馈函数可以定义为：

$$10 \quad r = r_{ACCURACY}(\eta r_{SATISFACTION} + \xi r_{EFFICIENCY}) \quad (8)$$

其中， η ， ξ 分别为切片需求满足度和链路效率的比例系数。 r 为反馈值，计算方式为需求满足度与 $r_{SATISFACTION}$ 和链路效率 $r_{EFFICIENCY}$ 加权求和，再与部署精度 $r_{ACCURACY}$ 相乘。由该公式可见，当切片需求满足度和链路效率越高，该切片部署策略的反馈值越高。反馈函数中的参数可以用神经网络进行拟合。

15 使用上述强化学习模块，并配合相应的反馈函数，即可根据获取到的当前网络状态，完成网络切片中每一个 VNF 和 PNF 的部署。

进一步的，对于数字孪生强化学习模块下多网络切片，可能会面临网络资源冲突的情况，此时需要进行策略调整。以下提供两种常见情况下的策略调整方式，在具体实施场景中，可以根据实际需要参考以下方式进行策略调整。

20 (1) 当网络切片请求在进行某个节点部署，由于之前的网络切片占用过多资源而无法找到资源满足需求的 PNF 时，根据网络切片的优先级排序，在不低于最低需求的条件下，按比例缩小前几个网络切片中每个 VNF 的资源，缩小的资源大小与网络切片的优先级成反比。缩小的资源大小 v 具体可以按如下公式计算。

$$25 \quad v = \frac{1}{1+e^{-\theta i}} \cdot (z - h) \quad (9)$$

其中， i 为切片优先级顺序， z 为之前被分配到的资源大小， h 为切片最低需求，参数 θ 可以在 $(0,1]$ 中取值。 θ 越大，资源缩减程度越大。优选方案中，当网络中切片需求较多时， θ 可以取 0.5~1。

(2) 当网络状态发生改变, 知识图谱模块中当前时刻的网络状态无法满足原时刻的切片需求时, 为了避免此资源冲突情况的发生, 在完成单个网络切片部署策略生成后, 就将部署策略传给知识图谱模块进行检测, 判断当前时刻的网络状态能否满足已经生成的网络切片决策。如果不满足, 则强化学习模块退回到不满足的切片步骤重新进行部署决策; 如果满足, 则更新网络状态, 基于更新后的网络状态继续进行下一次部署决策。

通过上述策略调整方式, 可以在决策失误时快速有效的进行调整, 确保下发的部署策略能够维持网络状态正常。

本实施例提供的基于强化学习的数字孪生网络切片的方法, 在进行节点选择和资源分配时引入了网络运行状态、资源利用率以及链路效率, 可以为网络切片生成合理的部署策略的同时保证网络运行状态。在多网络切片需求下, 根据网络切片的优先级和网络资源状态对网络切片的部署策略进行动态调整, 以此保证每个网络切片和整个网络能稳定运行。

实施例 2:

基于实施例 1 提供的基于强化学习的数字孪生网络切片的方法, 在不同的具体应用场景中, 在某些具体实施方式中, 可以通过本实施例中的具体实施方式对实施例 1 中提供的基于强化学习的数字孪生网络切片方法进行实现, 使用户的任务需求通过知识图谱模块转换成网络切片需求。

在某个具体实施场景中, 网络中存在 6 个 PNF, 它们的计算、存储、网络资源分别为 PNF_1 (100, 200, 95), PNF_2 (100, 60, 70), PNF_3 (200, 100, 80), PNF_4 (100, 200, 95), PNF_5 (200, 200, 200), PNF_6 (100, 100, 100), 其中 PNF_1 和 PNF_3 、 PNF_6 有连接关系, PNF_2 和 PNF_4 、 PNF_5 有连接关系, 物理可达。

用户发起两个切片请求: 网络切片 1 需要三个 VNF($VNF_1^1, VNF_2^1, VNF_3^1$), 每个 VNF 的计算、存储、网络最低需求分别为 VNF_1^1 : (10,10,10), VNF_2^1 : (10,20,10), VNF_3^1 : (20,10,20), 切片优先级为 1。网络切片 2 需要 4 个 VNF, 每个 VNF 的计算、存储、网络需求分别为 VNF_1^2 : (20,20,20), VNF_2^2 : (20,20,10), VNF_3^2 : (10,10,20), VNF_4^2 : (30,10,10), 切片优先级为 2。其中, VNF 的上标表示其所在的网络切片, 下标为每个网络切片中 VNF 自身的序号。

知识图谱模块根据前网络状态及可用资源构建当前网络知识图谱, 并将用

户的切片需求与网络知识图谱信息以三元组的形式传递给强化学习模块。其中，网络信息三元组为：(PNF₁, 连接, PNF₃), (PNF₁, 计算, 100), (PNF₁, 存储, 100), (PNF₁, 网络, 90), (PNF₂, 计算, 100), (PNF₂, 存储, 60), (PNF₂, 网络, 70), (PNF₃, 计算, 200), (PNF₃, 存储, 100), (PNF₃, 网络, 80), (PNF₄, 计算, 100), (PNF₄, 存储, 200), (PNF₄, 网络, 95), 其余三元组可依同样的规则建立。

网络切片 1 的一组切片需求三元组为：(切片 1, 优先级是, 1), (切片 1, 需要, VNF₁¹), (切片 1, 需要, VNF₂¹), (切片 1, 需要, VNF₃¹), (VNF₁¹, 计算, 10), (VNF₁¹, 存储, 10), (VNF₁¹, 网络, 10), (VNF₂¹, 计算, 10), (VNF₂¹, 存储, 20), (VNF₂¹, 网络, 10), (VNF₃¹, 计算, 20), (VNF₃¹, 存储, 10), (VNF₃¹, 网络, 20)。

网络切片 2 的一组切片需求三元组有 (切片 2, 优先级是, 2), (切片 2, 需要, VNF₁²), (切片 2, 需要, VNF₂²), (切片 2, 需要, VNF₃²), (切片 2, 需要, VNF₄²), (VNF₁², 计算, 20), (VNF₁², 存储, 20), (VNF₁², 网络, 20), (VNF₂², 计算, 20), (VNF₂², 存储, 20), (VNF₂², 网络, 10), (VNF₃², 计算, 10), (VNF₃², 存储, 10), (VNF₃², 网络, 20), (VNF₄², 计算, 30), (VNF₄², 存储, 10), (VNF₄², 网络, 10)。

按照步骤 101, 知识图谱将上述网络信息三元组和网络切片需求三元组, 以及网络信息和网络切片请求, 传到强化学习模块。强化学习模块依次分别先进行网络切片 1 和网络切片 2 的切片部署决策。

按照步骤 102, 基于知识图谱, 先为 VNF₁¹ 的节点部署构建节点状态矩阵 X 。矩阵的每一行为每个 PNF 节点的节点状态向量 $X_i = (X_i^1, X_i^2, X_i^3, X_i^4, X_i^5)$ 。其中, X_i^1 表示 PNF 节点 i 的可选性, 根据前几跳是否被选择过以及是否可达来判断, 可选为 1, 不可选为 0。 X_i^2 表示 PNF 节点 i 到上一个 PNF 节点的可用的最短链路的物理距离, 若为第一个节点, 则 X_i^2 都为 1。 X_i^3, X_i^4, X_i^5 分别表示 PNF 节点 i 上计算、存储和网络资源对需求的满足程度, 通过计算 PNF 节点 i 上的可用资源与切片对该节点的最小资源需求的差值得到。假设有 PNF₁: (100,100,90), PNF₂: (100,60,70), 对于第一个 VNF 都可选, 所以对应的节点状态向量分别为 $X_1 = [1,1,90,90,80,]$, $X_2 = [1,1,90,60,50]$ 。构建好节点状态矩阵后, 输入到节点选择概率模型中。节点概率选择模型是一个神经网络模型。模型首先对节点状态矩阵进行加权计算, 得到每一个 PNF 节点的评价分数, 然

后根据每个 PNF 的评价分数大小计算每个 PNF 节点被选择的概率。计算得到每个 PNF 的选择概率为 $[0.9, 0.7, \dots]$ ，其中 PNF_1 的概率最大为 0.9，所以选择 VNF_1^1 部署在 PNF_1 上。然后，构建资源评估向量为 $x_r = (10, 10, 10, 100, 100, 90)$ 输入到资源分配模型中计算需要分配的资源。资源分配模型通过函数将资源评估向量映射成资源分配向量，映射函数是事先通过神经网络拟合得到。计算得到输出为 $(85, 85, 80)$ ，则为 VNF_1^1 分配计算、存储、网络资源分别为 85、85、80。分配的资源是多少取决于资源分配模型，若模型训练得较好，得到的分配结果会更合理，后续不需要过多的调整；若模型训练的不够好，则无法得到最合理的分片，后续还需要进一步进行调整。

完成一个节点的部署动作后，环境知识图谱会进行更新，如 PNF_1 的可用资源减少为 $(15, 15, 10)$ 。接下来，根据前文提到的计算公式分别计算当前切片的部署精度、切片需求满足度以及链路效率，再根据反馈函数公式综合计算得到此次网络切片资源分配动作的奖励。其中，由于该节点部署成功，所以部署精度计算为 1。切片需求满足度，假设反馈系数为 2，根据节点需求和被分配的资源结合公式 (5) 可计算得 $2 \times (10/85 + 10/85 + 10/80)/3 = 0.24$ 。

该节点为第一个节点，所以链路效率计算为 1。假设切片需求满足度和链路效率的比例系数分别为 0.5，则根据公式 (7) 计算可得反馈奖励值为 $1 \times (0.5 \times 0.24 + 0.5 \times 1) = 0.62$ 。然后，结合反馈值利用梯度下降法更新一遍节点选择概率模型的参数。由于此次网络切片资源分配动作成功为一个网络切片分配资源并得到奖励，表明资源分配模型此次映射的方向是正确的，因此参数更新的结果是鼓励模型继续朝这个映射方向优化。

接下来根据第一个 VNF 节点选取的位置， PNF_1 ，来为 VNF_2^1 的节点部署构建节点状态矩阵 X 。假设有 $PNF_3(200, 100, 80)$ ， $PNF_4(100, 200, 95)$ ， PNF_3 到 PNF_1 的物理距离是 3， PNF_4 到 PNF_1 不可达，则对应的节点状态向量分别为 $X_3 = [1, 3, 190, 80, 70, \dots]$ ， $X_4 = [0, -1, 90, 180, 85]$ 。构建好节点状态矩阵后，输入到节点选择概率模型中计算得到每个 PNF 的选择概率为 $[\dots 0.8, 0.6, \dots]$ ，其中 PNF_3 的概率最大为 0.8，所以选择 VNF_2^1 部署在 PNF_3 上。然后，构建资源评估向量为 $x_r = (10, 20, 10, 200, 100, 80)$ 输入到资源分配模型中计算得到输出为 $(150, 80, 50)$ ，则为 VNF_2^1 分配计算、存储、网络资源分别为 150、80、50。

完成第二个节点的部署动作后，环境知识图谱进行更新，如 PNF_3 的可用资源减少为 $(50, 20, 30)$ 。再根据反馈函数计算奖励更新节点选择概率模型的参数。

计算方式如下：

$$w = w + \eta * \frac{\partial L(w)}{\partial w} * r$$

其中 w 为模型参数， η 为更新系数， r 为反馈函数值。

重复上述操作进行第三个节点 VNF_3^1 的部署，节点选择概率模型选择 VNF_3^1 部署在 PNF_6 上，资源分配模型为其分配计算、存储、网络资源分别为 60、50、60，即完成第一个网络切片请求的部署决策。完成网络切片 1 的切片部署决策后，强化学习模块开始进行第二个网络切片的部署决策。当为第二个网络切片的第一个节点 VNF_1^2 进行部署决策时发现只有 PNF_1 可选，但 PNF_1 由于在第一个网络切片时为 VNF_1^1 分配过多资源导致资源不够。此时利用资源分配调整公式 (9)，根据第一个网络切片的优先级计算需要缩小资源(42,42,40)。然后调整 VNF_1^1 的资源分配从(85,85,80)减少为(43,43,40)。现在 PNF_1 上的可用资源从(15,15,10)变为(57,57,50)，可以满足 VNF_1^2 的资源需求。于是继续按照先前步骤为第二个网络切片进行部署决策。

对第二个网络切片的每个节点进行部署决策后，这两个切片的部署决策为：
 S1: (PNF_1, PNF_3, PNF_5), VNF_1^1 : (57,57,50), VNF_2^1 : (150,80,50), VNF_3^1 : (60,70,100) ;
 S2: ($PNF_1, PNF_4, PNF_7, PNF_9$), VNF_1^2 : (20,20,20), VNF_2^2 : (30,30,25), VNF_3^2 : (80,80,100), VNF_4^2 : (50,20,20)。将部署决策传到知识图谱模块，知识图谱模块进行更新。

按照步骤 103，知识图谱将更新后的网络信息传回强化学习模块，从网络信息中获取到各 PNF 的失效平均时间和平均修复时间后根据公式 (1) 计算各 PNF 的可用性，假设计算得到的 PNF_1 的可用性为 0.99， PNF_2 的可用性为 0.98， PNF_3 的可用性为 0.96， PNF_4 的可用性为 0.97， PNF_6 的可用性为 0.98，设定的可用性阈值为 0.95，因此所有 PNF 都可用。再根据 PNF 的可用性以及 PNF 和 VNF 的资源条件可以得到切片的网络运行状态。其中，切片 S1 的 PNF 状态分别为 0000001111、0000101111、0101001111，切片 S1 的 VNF 初始状态分别为 0000011111、0101011111、1010011111；切片 S2 的 PNF 状态分别为 0000011111、0010111111、0111011111，切片 S2 的 VNF 初始状态分别为 0000111111、0101111111、1010111111、1110111111。于是网络状态评估为正常，然后将部署决策信息传至控制模块进行下发，完成用户的网络切片请求。

由上述具体实例可见，根据实施例 1 中提供的基于强化学习的数字孪生网络切片的方法，能够快速有效的完成数字孪生的网络切片，更好地保证网络资

源的高效利用，并确保网络动态变化的情况下能够保持网络状态正常。

实施例 3:

在上述实施例 1 至实施例 2 提供的基于强化学习的数字孪生网络切片的方法的基础上，本发明还提供了一种可用于实现上述方法的基于强化学习的数字孪生网络切片的装置，如图基于强化学习的数字孪生网络切片所示，是本发明
5 实施例的装置架构示意图。本实施例的基于强化学习的数字孪生网络切片的装置包括一个或多个处理器 11 以及存储器 12。其中，图 6 中以一个处理器 11 为例。

10 处理器 11 和存储器 12 可以通过总线或者其他方式连接，图 6 中以通过总线连接为例。

存储器 12 作为一种基于强化学习的数字孪生网络切片的方法非易失性计算机可读存储介质，可用于存储非易失性软件程序、非易失性计算机可执行程序以及模块，如实施例 1 至实施例 2 中的基于强化学习的数字孪生网络切片的方法。处理器 11 通过运行存储在存储器 12 中的非易失性软件程序、指令以及
15 模块，从而执行基于强化学习的数字孪生网络切片的装置的各种功能应用以及数据处理，即实现实施例 1 至实施例 2 的基于强化学习的数字孪生网络切片的方法。

存储器 12 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
20 实施例中，存储器 12 可选包括相对于处理器 11 远程设置的存储器，这些远程存储器可以通过网络连接至处理器 11。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

程序指令/模块存储在存储器 12 中，当被一个或者多个处理器 11 执行时，执行上述实施例 1 至实施例 2 中的基于强化学习的数字孪生网络切片的方法，
25 例如，执行以上描述的图 2 和图 3 所示的各个步骤。

本领域普通技术人员可以理解实施例的各种方法中的全部或部分步骤是可以通过程序来指令相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：只读存储器（Read Only Memory，简称为：ROM）、
随机存取存储器（Random Access Memory，简称为：RAM）、磁盘或光盘等。

30 以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等，均应包含在本发明的保护范围之内。

权利要求书

1.一种基于强化学习的数字孪生网络切片的方法，其特征在于，具体包括：

强化学习模块获取知识图谱中当前网络信息三元组和网络切片需求三元组，每个网络切片对应一组网络切片需求三元组，其中，当前网络信息三元组包括网络中所有 PNF 连接关系和每个 PNF 的资源属性，网络切片需求三元组包括切片需求的所有 VNF 连接关系、每个 VNF 的资源需求和切片优先级；

强化学习模块生成至少一个网络切片的决策动作，每个决策动作包含根据节点选择概率模型和知识图谱查找 PNF 节点以及根据资源分配模型对每个 PNF 节点分配给 VNF 的资源映射为资源分配向量，知识图谱根据决策动作进行决策以获得决策结果，决策结果为网络切片 VNF 与网络切片 PNF 的部署位置；

知识图谱根据决策结果进行更新，并将更新后的网络状态传回给强化学习模块，强化学习模块确认网络状态正常后，下发网络切片的部署策略。

2.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块生成至少一个网络切片的决策动作，具体包括：

根据切片需求，基于知识图谱，找到节点选择概率模型中概率最大的 PNF 作为第一个节点；

基于知识图谱得到资源评估向量，输入到资源分配模型中为第一 VNF 节点分配资源，并更新知识图谱；

根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到可能的下一个 VNF 节点的 PNF 位置选择，依次查找直至所有节点都部署分配完毕。

3.根据权利要求 2 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述找到节点选择概率模型中概率最大的 PNF 作为第一个节点，具体包括：

找到第一个 VNF 节点的动作空间中所有可能的 PNF 位置选择作为第一 VNF 节点的部署动作空间，并得到节点状态矩阵；

将节点状态矩阵输入到节点选择概率模型中，计算每个节点被选择的概率，选择其中概率最大的 PNF 作为第一个节点；

基于知识图谱得到第一个节点的资源评估向量，将评估向量输入到资源分配模型中，为第一 VNF 节点分配资源。

4.根据权利要求 2 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述找到可能的下一个 VNF 节点的 PNF 位置选择，具体包括：

根据上一个 VNF 节点选取的位置，基于知识图谱里包含的 PNF 连接关系进行知识推理，找到所有可能的下一个 VNF 节点的 PNF 位置选择，作为下一个 VNF 节点的部署动作空间；

基于知识图谱得到节点状态矩阵，依次向后进行下一个节点的部署和资源分配，直至所有节点都部署分配完毕，以所有节点的分配结果完成一整个切片部署策略的选择。

5.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述知识图谱根据决策结果进行更新，具体包括：

知识图谱根据强化学习的资源分配策略进行更新，并根据反馈函数计算反馈值；结合反馈值利用梯度下降法更新节点选择概率模型的参数。

6.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述节点选择概率模型具体包括：

将 PNF 之间的距离、切片对于 VNF 节点的资源需求和 PNF 上的资源状态映射成可能选择的概率，以便于选择强化学习模块获取概率最大的 PNF 节点进行部署。

7.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述资源分配模型具体包括：

通过函数将资源评估向量映射成资源分配向量，以便于表示 PNF 节点分配给 VNF 的计算、存储、网络资源。

8.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块确认网络状态正常，还包括：

当网络切片请求在进行某个节点部署，由于之前的网络切片占用过多资源而无法找到资源满足需求的 PNF 时，根据网络切片的优先级排序，在不低于最低需求的条件下，按比例缩小前几个网络切片中每个 VNF 的资源，缩小的资源大小与网络切片的优先级成反比。

9.根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块确认网络状态正常，还包括：

当网络状态发生改变，知识图谱模块中当前时刻的网络状态无法满足原时刻的切片需求时，则强化学习模块退回到不满足的切片步骤重新进行部署决策。

10. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法包括：

- 5 所述网络信息三元组中，第一个实体表示 PNF，关系表示该三元组的资源类型，第二个实体表示第一个实体所需的该类型资源的值；

所述网络切片需求三元组中，第一个实体表示网络切片或 VNF，关系表示该三元组的资源类型，第二个实体表示第一个实体所需的该类型资源的值。

- 10 11. 根据权利要求 10 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法还包括：

所述网络切片需求三元组中，当第一个实体为网络切片时，关系的类型为所需的 VNF 或优先级，第二个实体为第一个实体需要的 VNF 或优先级的值。

- 15 12. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述知识图谱根据决策结果进行更新，并将更新后的网络状态传回给强化学习模块，具体包括：

将一个网络切片需求转化为至少一个 VNF 的分配需求，对于每个 VNF 的分配过程，确认分配在哪个 PNF 上，并根据需求分配相应的资源；

每分配完成一个 VNF 后，对知识图谱中的网络信息三元组进行更新，以便根据更新后的网络信息三元组对下一个 VNF 进行预测。

- 20 13. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块确认网络状态正常后，下发网络切片的部署策略，具体包括：

知识图谱将更新后的网络状态传回强化学习模块，强化学习模块对调整后的网络状态做出评估，确认使用该切片策略后网络状态正常后，将部署决策信息传
25 至控制模块进行下发。

14. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法还包括：

完成所有网络切片部署决策后，将部署决策传到知识图谱模块，对知识图谱模块进行更新；

知识图谱将更新后的网络状态传回强化学习模块，以便使知识图谱中的对应关系能够在网络环境出现动态变化时与当前环境相符。

15. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块具体包括：

5 使用以下数学模型对网络运行状态进行描述：

$N^P = \{n_1^P, n_2^P, \dots\}$ 表示物理集群层中的 PNF；

$l^P = \{l_{1,2}^P, l_{1,3}^P, \dots\}$ 表示物理集群层中的 PNF 与 PNF 之间的物理链路；

$S = \{S_1, S_2, \dots, S_b\}$ 表示网络中存在的网络切片用，总计有 b 个网络切片；

在一个网络切片 S_s 中， $N^S = \{n_1^S, n_2^S, \dots\}$ 表示 VNF 节点， $l_s = \{l_{1,2}^S, l_{1,3}^S, \dots\}$ 表示

10 VNF 节点之间的虚拟链路；

对于 PNF n_i^P 的计算、存储、网络资源分别表示为 $H_{n_i^P}^1, H_{n_i^P}^2, H_{n_i^P}^3$ ；

切片中 VNF 节点 n_i^S 的计算、存储、网络资源分别表示为 $h_{n_i^S}^1, h_{n_i^S}^2, h_{n_i^S}^3$ 。

16. 根据权利要求 15 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块还包括：

15 每个 PNF 或 VNF 可靠性由失效平均时间衡量，可服务性由平均修复时间衡量；PNF (VNF) 的可用性 A_{PNF} (A_{VNF}) 使用以下公式表示：

$$A_{PNF} = \frac{MTTF}{MTTF + MTTR};$$

其中，整个切片的可用性等于构成切片 PNF 和 VNF 的可用性最小值，用于评估网络运行状态。

20 17. 根据权利要求 15 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块还包括：

VNF 和 PNF 的位置关系用 $a_{n_i^S n_i^P}$ 使用以下公式表示：

$$a_{n_i^S n_i^P} = \begin{cases} 1 & \text{if } n_i^S \text{ 在 } n_i^P \text{ 上;} \\ 0 & \text{otherwise} \end{cases}$$

其中，如果 VNF n_i^S 在 PNF n_i^P 上，则 $a_{n_i^S n_i^P} = 1$ ，否则为 0。

25 18. 根据权利要求 15 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述强化学习模块还包括：

VNF n_i^S 和 PNF n_i^P 的计算、存储和网络资源满足以下条件：

$$\forall q=1,2,3 \sum_{s=1}^b h_{n_i^S}^q a_{n_i^S n_i^P} < H_{n_i^P}^q;$$

其中，根据 PNF、VNF 可用性以及资源条件定义网络运行态，对网络切片时，选择 VNF 并选择所部署的 PNF，按照空间位置关系对 PNF 排序。

19. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法还包括：

5 节点概率选择模型对节点状态矩阵进行加权计算： $Y = WX + b$ ，以得到每一个 PNF 节点是否可以作为下一个 PNF 节点的评价分数 $Y = (y_1, y_2, \dots, y_n)$ ；

根据评价分数计算每个 PNF 节点作为被选择的概率 $P_i = \frac{y_i}{\sum_{k=0}^N y_k}$ ；

其中，W 为神经元权重参数，b 为模型偏置参数，均由事先的模型训练得到。

10 20. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法还包括：

奖励更新策略模型的反馈函数使用以下公式表示：

$$r = r_{\text{ACCURACY}}(\eta r_{\text{SATISFACATION}} + \xi r_{\text{EFFICIENCY}});$$

15 其中， η ， ξ 分别为切片需求满足度和链路效率的比例系数，r 为反馈值，计算方式为需求满足度与 $r_{\text{SATISFACATION}}$ 和链路效率 $r_{\text{EFFICIENCY}}$ 加权求和，再与部署精度 r_{ACCURACY} 相乘。

21. 根据权利要求 1 所述的基于强化学习的数字孪生网络切片的方法，其特征在于，所述方法还包括：

当网络切片请求在过多资源而无法找到资源满足需求的 PNF 时，缩小的资源大小 v 使用如下公式计算：

$$20 \quad v = \frac{1}{1+e^{-\theta i}} \cdot (z - h);$$

其中， i 为切片优先级顺序， z 为之前被分配到的资源大小， h 为切片最低需求，参数 θ 可以在 $(0,1]$ 中取值。 θ 越大，资源缩减程度越大。

22. 一种基于强化学习的数字孪生网络切片的装置，其特征在于：

25 包括至少一个处理器和存储器，所述至少一个处理器和存储器之间通过数据总线连接，所述存储器存储能被所述至少一个处理器执行的指令，所述指令在被所述处理器执行后，用于完成权利要求 1-21 中任一项所述的基于强化学习的数字孪生网络切片的方法。

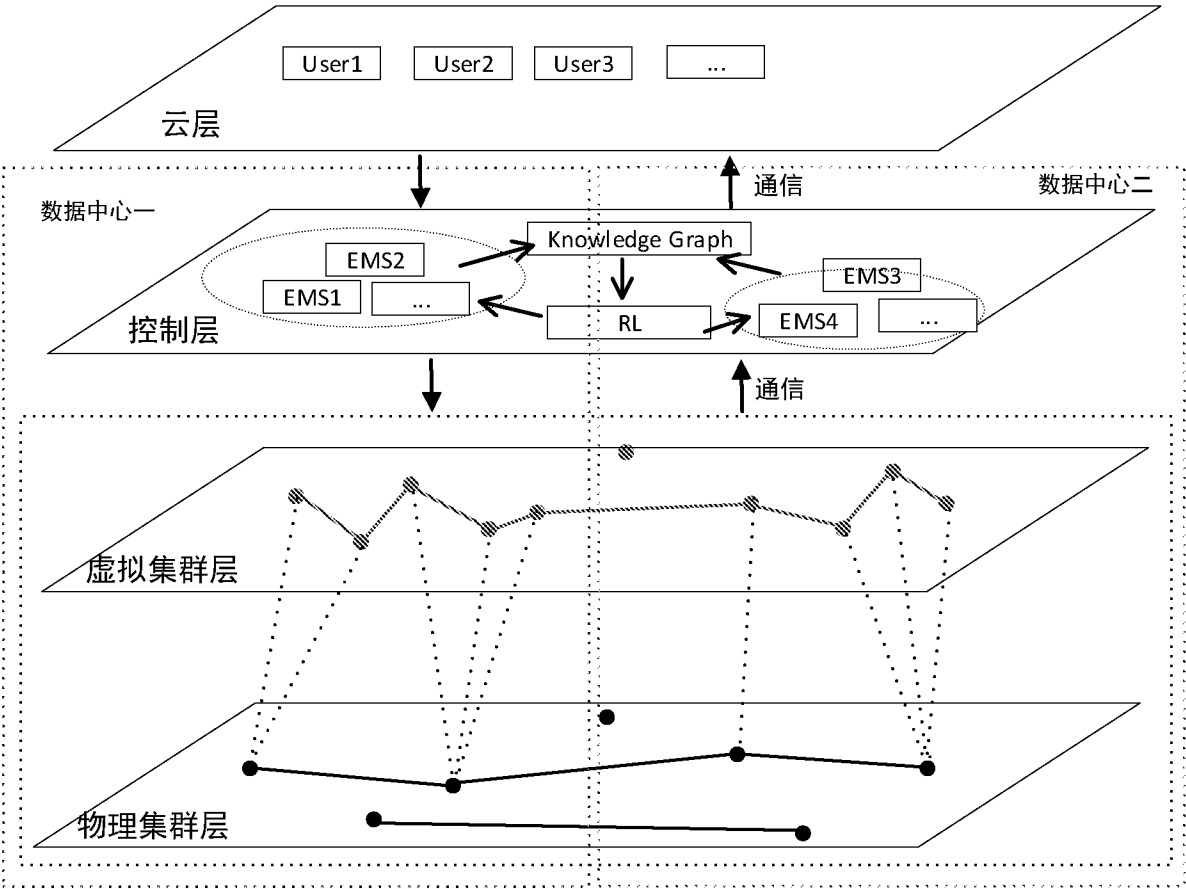


图 1

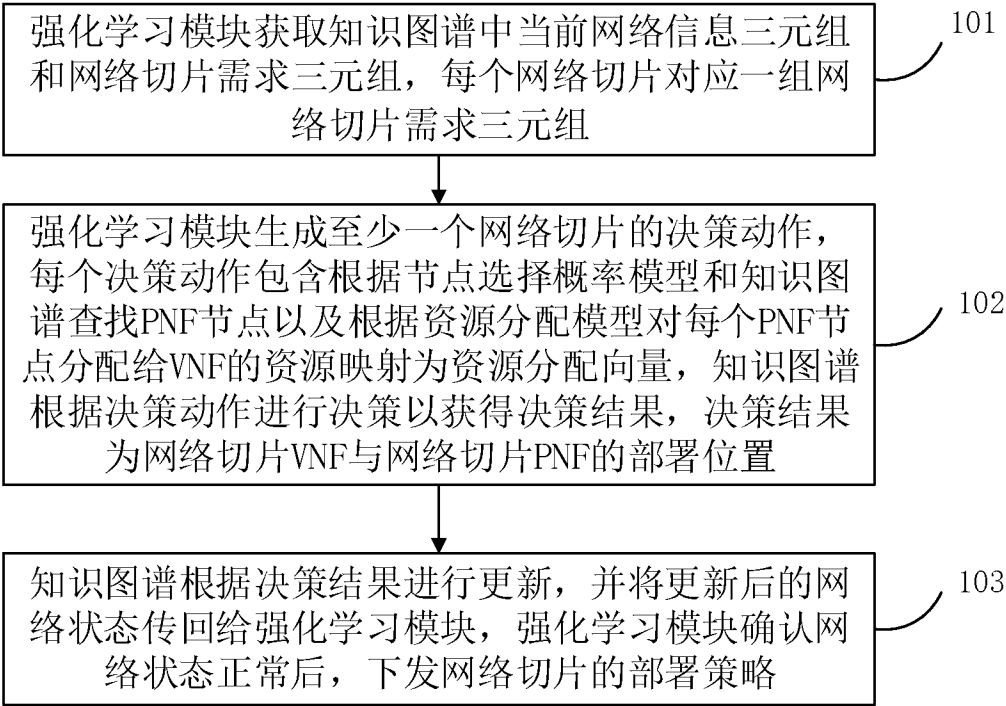


图 2

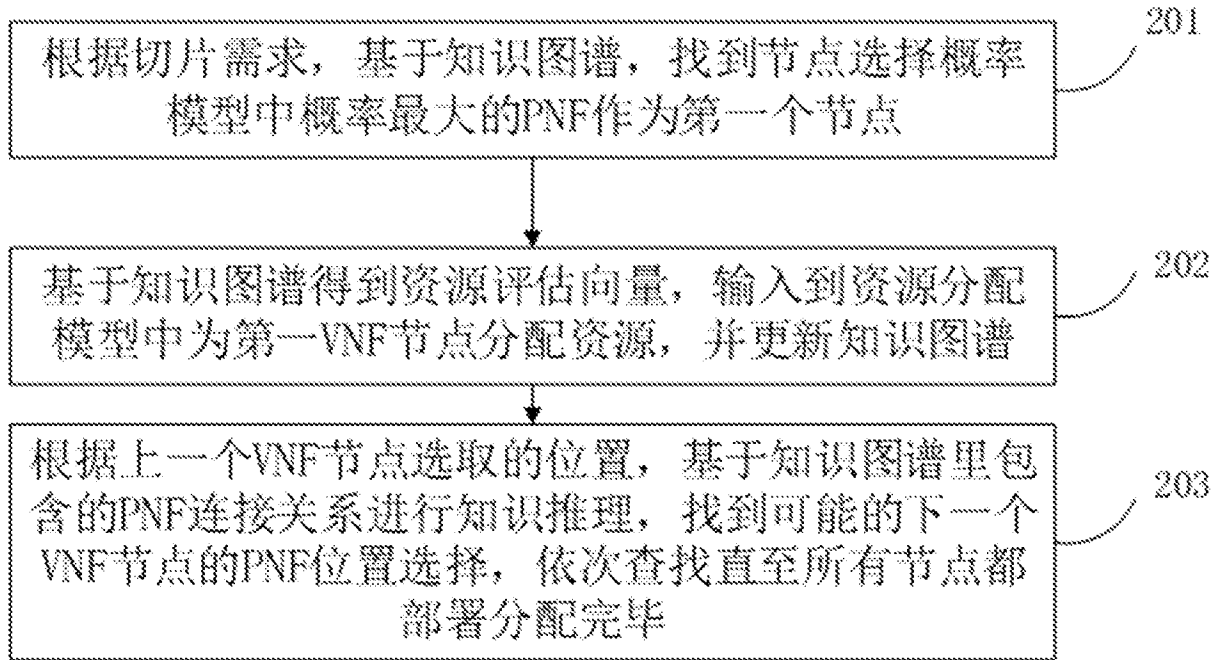


图 3

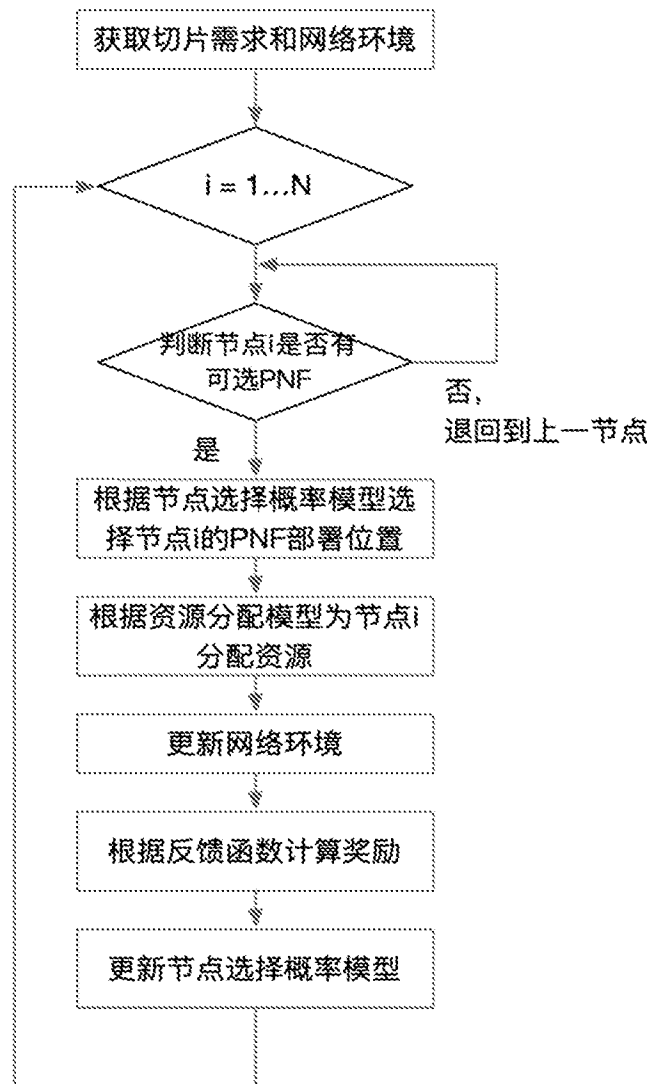


图 4

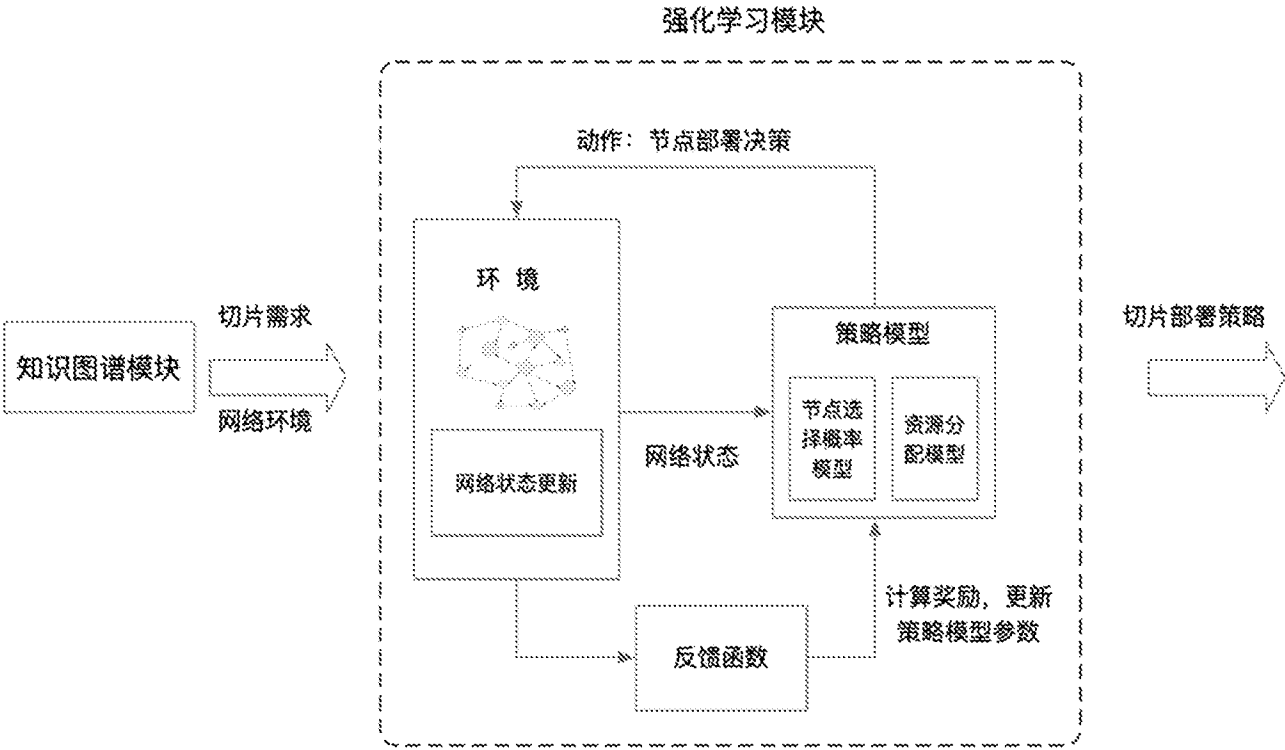


图 5

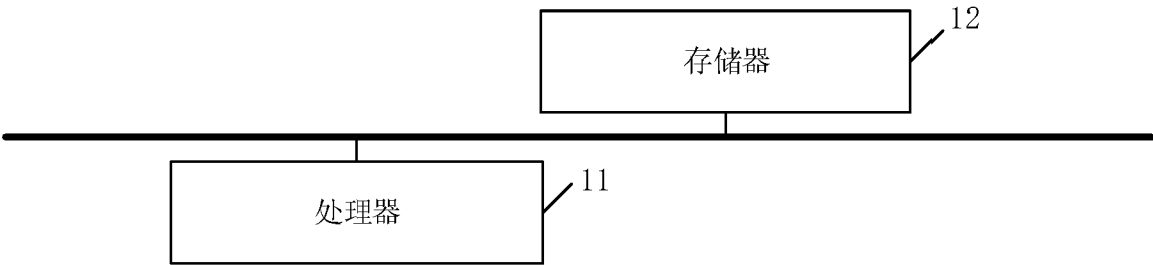


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/070341

A. CLASSIFICATION OF SUBJECT MATTER

H04L 41/0894(2022.01)i; G06F 16/36(2019.01)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)

IPC: H04L G06F

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)

CNTXT, ENTXT, DWPI: 网络切片, 资源, 分配, 孪生, 强化学习, 知识图谱, 物理网络功能, 虚拟网络功能, 更新, network slice, resource, allocation, twinning, reinforcement learning, knowledge graph, PNF, VNF, update

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116389266 A (FIBERHOME TELECOMMUNICATION TECHNOLOGIES CO., LTD.) 04 July 2023 (2023-07-04) claims 1-10, and description, paragraphs 0028-0091	1-22
A	CN 114666220 A (CHINA TELECOM CORPORATION LIMITED) 24 June 2022 (2022-06-24) description, paragraphs 0085-0188	1-22
A	CN 113098714 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 09 July 2021 (2021-07-09) entire document	1-22
A	CN 115603850 A (ZTE CORP.) 13 January 2023 (2023-01-13) entire document	1-22
A	US 2021345357 A1 (KONINKLIJKE KPN N.V. et al.) 04 November 2021 (2021-11-04) entire document	1-22



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

“D” document cited by the applicant in the international application

“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

18 March 2024

Date of mailing of the international search report

21 March 2024

Name and mailing address of the ISA/CN

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

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/070341

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116389266	A	04 July 2023	None			
CN	114666220	A	24 June 2022	None			
CN	113098714	A	09 July 2021	None			
CN	115603850	A	13 January 2023	WO	2023273940	A1	05 January 2023
US	2021345357	A1	04 November 2021	WO	2020074687	A1	16 April 2020
				CN	112789832	A	11 May 2021
				EP	3864804	A1	18 August 2021
				US	2023007662	A1	05 January 2023

A. 主题的分类

H04L 41/0894(2022.01)i; G06F 16/36(2019.01)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)

IPC: H04L G06F

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))

CNTEXT, ENTXT, DWPI: 网络切片, 资源, 分配, 孪生, 强化学习, 知识图谱, 物理网络功能, 虚拟网络功能, 更新, network slice, resource, allocation, twinning, reinforcement learning, knowledge graph, PNF, VNF, update

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
PX	CN 116389266 A (烽火通信科技股份有限公司) 2023年7月4日 (2023 - 07 - 04) 权利要求1-10, 说明书第0028-0091段	1-22
A	CN 114666220 A (中国电信股份有限公司) 2022年6月24日 (2022 - 06 - 24) 说明书第0085-0188段	1-22
A	CN 113098714 A (南京邮电大学) 2021年7月9日 (2021 - 07 - 09) 全文	1-22
A	CN 115603850 A (中兴通讯股份有限公司) 2023年1月13日 (2023 - 01 - 13) 全文	1-22
A	US 2021345357 A1 (KONINKLIJKE KPN N.V.等) 2021年11月4日 (2021 - 11 - 04) 全文	1-22

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2024年3月18日

国际检索报告邮寄日期

2024年3月21日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

王国纲

电话号码 (+86) 010-53961755

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/070341

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	116389266	A	2023年7月4日	无			
CN	114666220	A	2022年6月24日	无			
CN	113098714	A	2021年7月9日	无			
CN	115603850	A	2023年1月13日	WO	2023273940	A1	2023年1月5日
US	2021345357	A1	2021年11月4日	WO	2020074687	A1	2020年4月16日
				CN	112789832	A	2021年5月11日
				EP	3864804	A1	2021年8月18日
				US	2023007662	A1	2023年1月5日