

一种任务求解方法及其装置

本申请要求于 2021 年 9 月 30 日提交中国专利局、申请号为 202111166727.3、发明名称为“一种任务求解方法及其装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及调度领域，尤其涉及一种任务求解方法及其装置。

背景技术

10 调度问题是大型制造、物流、生产等环节中最常见的问题之一，在不同的场景下，调度总是有不同的意义。例如：物流调度主要是指在物流过程中，物流公司根据待发货物的重量、去向、规格、加急程度等对所属的车辆和人员进行合理的安排和调度；而生产环境中的调度是根据不同产线中不同机器的产能以及生产需求，在若干任务(job)中完成对任务的排序以及任务和机器（可调度资源）之间的匹配；大型制造工厂/机场的工人/空乘排班
15 (timetabling) 也是调度问题的一种，这是由于这类问题的目标也是依照工人/空乘的工作特点以及场景需要在不同的时间段内完成最优匹配。因此，核心是排序以及最优分配，而不局限任务是人还是货物。一般来讲，调度问题的目标是在给定任务数的前提下得到最小总工时(makespan)所对应的排序。

很多的调度问题（如排产、线体调度、加工网络布局等）都可以建模成一个数学问题
20 来求解，线性规划(linear programming, LP)是其中使用最广的一类建模方法。线性规划模型可以包括目标函数和约束条件，其中，目标函数是指根据待优化的目标和影响该目标的变量所设计的函数。例如，在排产问题中，整个排产的目标通常是在满足所有资源约束的情况下，找出一个最好的加工计划，使得需求的满足率最高，同时整体的成本最小（例如成本可以包括但不限于加工成本、库存成本、转运成本），此时，该目标函数可以是用于
25 表示满足率最大化以及成本最小化的函数。另外，约束条件是指在求解目标函数的过程中所要满足的其他限制条件。

然而，随着场景复杂性的提升，线性规划任务中规划约束的数量通常较多，求解线性规划任务所需的求解时间很长。

30

发明内容

第一方面，本申请提供了一种任务求解方法，所述方法包括：

获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；

在一种可能的实现中，终端设备可以将第一线性规划任务作为待求解模型传递至服务器，进而服务器可以获取第一线性规划任务。又例如，终端设备可以将第一线性规划任务
35 作为求解待求解模型时所使用的的先验信息，也就是将包括第一线性规划任务在内的至少一个历史模型传递至服务器，进而服务器可以获取包括第一线性规划任务在内的至少一个

历史模型；

通常情况下，线性规划任务中约束的数量较多时，求解线性规划任务所需的求解时间很长，因此可以首先选择线性规划任务中的部分约束，并基于选择的部分约束进行求解，并将求解结果（选择的部分约束中求解变量的状态）赋予线性规划任务，相当于将求解结果作为线性规划任务的初始值，并对线性规划任务进行求解，如果选择的部分约束的求解结果与求解线性规划任务后的求解结果基本一致（或者描述为较为接近，这里的求解结果基本一致可以理解为相同的约束中的求解变量的参数值基本一致），则求解线性规划任务时所需的迭代次数较少，也就是可以提高线性规划任务的求解速度。

因此，在求解第一线性规划任务之前，需要选择一部分约束进行求解，且得到的求解结果与第一线性规划任务的求解结果较为接近。

本申请实施例中，基于重要性来作为从多个第一规划约束中选择部分约束的依据，其中重要性可以指示对于降低所述第一线性规划任务的求解时间的贡献程度。其中，所谓指示对于降低所述第一线性规划任务的求解时间的贡献程度，可以理解为，在构建待求解模型的子模型时（也就是从待求解模型中的约束中选择一部分来构建子模型），在将子模型求解的结果作为待求解模型的初始值的情况下，子模型包括该第一线性规划任务对于降低求解待求解模型的求解时间的贡献程度。

具体的，可以将重要性作为采样概率对多个第一规划约束进行采样可以得到一个子任务（第二规划任务），该子任务中的规划约束为对于降低所述第一线性规划任务的求解时间的贡献程度较高的约束，因此，对第二规划任务求解后得到的求解结果与对第二规划任务求解后得到的求解结果基本一致（或者描述为较为接近，这里的求解结果基本一致可以理解为相同的约束中的求解变量的参数值基本一致）。

在一种可能的实现中，该重要性可以基于先验信息得到，其中先验信息可以为多个和待求解模型（第一线性规划任务）同类的一个或多个模型，所谓同类，可以理解为线性规划任务包括的约束类型、数量一致或者基本一致，也就是说第一线性规划任务和先验信息中的一个或多个模型包括的约束类型、数量一致或者基本一致。

在确定各个约束的重要性时，可以直接对先验信息中各个线性规划任务进行求解，得到求解各个线性规划任务的求解时间（若线性规划任务已经被求解过，也可以直接获取到求解时间），然后对先验信息中的一个或多个线性规划模型进行子模型的提取、求解，并基于求解结果对线性规划模型进行初始化，然后对线性规划模型进行求解，以得到另一个求解时间，通过比对直接求解各个线性规划任务的求解时间，和对基于子模型求解结果进行初始化后的线性规划任务的求解时间，就可以知晓子模型中的约束对于降低线性规划任务的求解时间的贡献程度。

其中，为了能够量化上述重要性，可以对待求解模型的各个约束赋予一个概率（该概率可以为上述重要性的量化），并基于概率对待求解模型进行采样，以构建子模型，并基于比对直接求解各个线性规划任务的求解时间，和对基于子模型求解结果进行初始化后的线性规划任务的求解时间，来更新概率，迭代一定的次数，就可以得到用于准确描述约束对于降低线性规划任务的求解时间的重要性信息（关于重要性的更新方式将在候选实施例中

描述)。

基于上述重要性采样得到的子模型(例如本申请实施例中的第二线性规划任务)中的各个规划约束对降低线性规划任务的求解时间的贡献程度很高,因此,基于第二线性规划任务的求解结果初始化后的第一线性规划模型的求解过程的时间可以被大大缩短;

- 5 根据所述重要性,对所述多个第一规划约束进行采样,以得到所述获取所述多个第一规划约束的子集,其中,所述重要性用于确定第一规划约束的采样概率;

根据所述多个第一规划约束的子集构建第二线性规划任务;

- 10 对所述第二线性规划任务进行求解,得到第一求解结果;将所述第一求解结果作为所述第一线性规划任务的初始值,并对初始化后的所述第一线性规划任务进行求解,得到第二求解结果。

可选的,第一求解结果可以包括多个第一规划约束的子集中的至少一个求解变量的参数值(或者参数状态);

- 15 本申请实施例提供了一种任务求解方法,所述方法包括:获取第一线性规划任务,所述第一线性规划任务包括多个第一规划约束;获取所述多个第一规划约束中每个第一规划约束的重要性,所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度;根据所述重要性,对所述多个第一规划约束进行采样,以得到所述获取所述多个第一规划约束的子集,其中,所述重要性用于确定第一规划约束的采样概率;根据所述多个第一规划约束的子集构建第二线性规划任务;将所述第一求解结果作为所述第一线性规划任务的初始值,并对初始化后的所述第一线性规划任务进行求解,得到第二求解结果。通过上述方式,一方面选取第一线性规划任务中的关键约束构建子模型(第二线性规划任务),然后利用子模型的解来第一线性规划任务从而加速求解过程。另一方面,在构建子模型时,基于重要性对第一线性规划任务进行采样,由于重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度,使采样得到的子模型的解接近于最优解,从而加速第一线性规划任务的求解过程。

25

在一种可能的实现中,第一线性规划模型可以为用于更新重要性所使用的先验信息,在这种情况下,可以基于比对直接求解各个线性规划任务的求解时间,和对基于子模型求解结果进行初始化后的线性规划任务的求解时间,来更新概率,迭代一定的次数,就可以得到用于准确描述约束对于降低线性规划任务的求解时间的重要性信息。

- 30 具体的,可以获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间,所述第一求解时间和所述第三求解时间的加和可以作为上述基于子模型求解结果进行初始化后的线性规划任务的求解时间,并对所述第一线性规划任务进行求解,得到求解所述第一线性规划任务的第三求解时间,其中,第三求解时间可以认为是直接求解第一线性规划任务的求解时间,根据所述第一求解时间和所述第三求解时间的加和,相对于所述第一求解时间的减少程度,就可以更新所述每个第一规划约束的重要性,其中,更新后的所述重要性与所述减少程度正相关。

35

应理解,还可以将采样得到第二线性规划任务的过程所需的时间也作为基于子模型求

-4-

解结果进行初始化后的线性规划任务的求解时间的一部分，也就是可以获取对所述多个第一规划约束进行采样的采样时间，并根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性。

5

在一种可能的实现中，所述方法还包括：获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；对所述第四线性规划任务进行求解，得到第三求解结果；将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三线性规划任务进行求解，得到第四求解结果。

10
15

在一种可能的实现中，所述第二线性规划任务包括M个求解变量，所述第一求解结果包括各个求解变量的参数值；

所述将所述第一求解结果作为所述第一线性规划任务的初始值，包括：

将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的M个求解变量的参数值。

20

在一种可能的实现中，所述第一线性规划任务用于为至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

25

在一种可能的实现中，所述得到第二求解结果之后，可能还会新加约束（例如第二规划约束），可以获取第二规划约束，所述第二规划约束包括松弛变量以及所述松弛变量的上界，将所述第二规划约束增加至所述第一线性规划任务，以得到更新后的所述第一线性规划任务，并求解所述更新后的所述第一线性规划任务。具体可以是给新添加的约束中涉及到的新变量（松弛变量）根据其实际含义设置一个合理的bound（数值范围），并修改这些变量的初始状态值，将状态改到其设置的bound上，使得初始解对于新添加的约束是可行的，进而加速线性规划的求解。

30

本申请可以在连续求解场景中针对新添加的松弛变量，预先设置变量上界和初始状态，从而加速连续求解。

35

本申请实施例还提供一个系统，其中，系统可以包括终端设备以及服务器，其中，终端设备可以发送待求解模型（第一线性规划任务）以及包括多个历史模型在内的先验信息至服务器，服务器可以基于包括多个历史模型在内的先验信息计算得到收敛的重要性，并

-5-

根据重要性采样第一线性规划任务，并执行上述实施例中步骤301至步骤303的步骤，以得到第二求解结果，并将第二求解结果回传至终端设备。

此外终端设备可以发送待求解模型以及包括多个历史模型（包括第一线性规划任务）在内的先验信息至服务器，服务器可以基于包括多个历史模型在内的先验信息计算得到收敛的重要性（计算过程可以参照上述实施例中关于更新重要性相关的描述），并根据重要性采样待求解模型，以及求解待求解模型。

第二方面，本申请提供了一种任务求解装置，所述装置包括：

10 获取模块，用于获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；

获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度；

采样模块，用于根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；

15 根据所述多个第一规划约束的子集构建第二线性规划任务；

求解模块，用于对所述第二线性规划任务进行求解，得到第一求解结果；

将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。

20 在一种可能的实现中，所述第一线性规划任务和所述第二线性规划任务均包括规划目标。

在一种可能的实现中，所述获取模块，还用于：

25 获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间；

所述求解模块，还用于对所述第一线性规划任务进行求解，得到求解所述第一线性规划任务的第三求解时间；

所述装置还包括：

30 重要性更新模块，用于根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，其中，更新后的所述重要性与所述减少程度正相关。

在一种可能的实现中，所述获取模块，还用于：

获取对所述多个第一规划约束进行采样的采样时间；

35 所述重要性更新模块，具体用于：

根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性。

在一种可能的实现中，所述获取模块，还用于：

获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；

- 5 获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；

所述采样模块，还用于：

- 10 根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；

所述求解模块，还用于：

对所述第四线性规划任务进行求解，得到第三求解结果；

将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三线性规划任务进行求解，得到第四求解结果。

15

在一种可能的实现中，所述第二线性规划任务包括M个求解变量，所述第一求解结果包括各个求解变量的参数值；

所述求解模块，具体用于：

- 20 将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的M个求解变量的参数值。

在一种可能的实现中，所述第一线性规划任务用于为至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

25

在一种可能的实现中，所述获取模块，还用于：

在所述得到第二求解结果之后，获取第二规划约束，所述第二规划约束包括松弛变量以及所述松弛变量的上界；

- 30 所述求解模块，还用于将所述第二规划约束增加至所述第一线性规划任务，以得到更新后的所述第一线性规划任务，并求解所述更新后的第一线性规划任务。

第三方面，本申请实施例提供了一种装置，包括存储器、处理器以及总线系统，其中，存储器用于存储程序，处理器用于执行存储器中的程序，以执行如上述第一方面及第一方面任一可选的方法。

- 35 第四方面，本发明实施例还提供一种系统，该系统包括至少一个处理器，至少一个存储器以及至少一个通信接口；处理器、存储器和通信接口通过通信总线连接并完成相互间的通信；

存储器用于存储执行以上方案的应用程序代码，并由处理器来控制执行。所述处理器用于执行所述存储器中存储的应用程序代码，以得到任务调度结果；其中存储器存储的代码可执行以上提供的一种任务求解方法。

通信接口，用于与其他设备或通信网络通信，以将所述任务求解结果发送至所述设备或通信网络。

第五方面，本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面及其任一可选的方法。

第六方面，本申请实施例提供了一种计算机可读存储介质，所述计算机存储介质存储有一个或多个指令，所述指令在由一个或多个计算机执行时使得所述一个或多个计算机实施上述第二方面及其任一可选的系统。

第七方面，本申请实施例提供了一种计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面及其任一可选的方法。

第八方面，本申请提供了一种芯片系统，该芯片系统包括处理器，用于支持终端设备或服务器实现上述方面中所涉及的功能，例如，发送或处理上述方法中所涉及的数据；或，信息。在一种可能的设计中，所述芯片系统还包括存储器，所述存储器，用于保存终端设备或服务器必要的程序指令和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

本申请实施例提供了一种任务求解方法，所述方法包括：获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度；根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；根据所述多个第一规划约束的子集构建第二线性规划任务；将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。通过上述方式，一方面选取第一线性规划任务中的关键约束构建子模型（第二线性规划任务），然后利用子模型的解来第一线性规划任务从而加速求解过程。另一方面，在构建子模型时，基于重要性对第一线性规划任务进行采样，由于重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度，使采样得到的子模型的解接近于最优解，从而加速第一线性规划任务的求解过程。

附图说明

- 图 1 为本申请实施例提供一种应用架构示意图；
- 图 2 为本申请实施例提供一种服务器的架构示意图；
- 图 3 为本申请实施例提供一种任务求解方法的流程示意图；
- 图 4 为本申请实施例提供一种任务求解方法的流程示意图；
- 图 5 为本申请实施例提供一种排产任务示意图；

图 6 为本实施例提供的一种任务求解装置的结构示意；
图 7 为本申请实施例提供的终端设备的一种结构示意图；
图 8 为本申请实施例提供的服务器一种结构示意图；
图 9 为本申请实施例提供的芯片的一种结构示意图。

5

具体实施方式

下面结合本发明实施例中的附图对本发明实施例进行描述。本发明的实施方式部分使用的术语仅用于对本发明的具体实施例进行解释，而非旨在限定本发明。

10 下面结合附图，对本申请的实施例进行描述。本领域普通技术人员可知，随着技术的发展和场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，这仅仅是描述本申请的实施例中对相同属性的对象在描述时所采用的区分方式。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便
15 包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

本申请实施例可以应用于多种场景（比如供应链、云计算、调度、存储优化等）的线性规划优化问题的求解中，加速线性规划求解器求解这些问题的效率。

20 参照图 1，图 1 为本申请实施例提供的应用结构示意，本申请提供的任务求解方法可以作为求解器部署在云侧的服务器，终端设备可以将待求解模型（例如线性规划任务）传递至云侧的服务器，云侧的服务器可以基于自身部署的求解器对待求解模型进行求解，并将求解结果传递至终端设备。

比如用户可以根据自己的业务场景构建待求解模型，在求解时候，可以将一部分历史上同类问题的模型传递至服务器，服务器可以调用求解器较快的输出用户输入模型的最优解。用户可以根据这个解去使用平台提供的功能生成数据报表或者自己对它进行处理得到想要的结果。

参照图 2，图 2 为本申请实施例提供的服务器的架构示意。具体的，服务器 200 由一个或多个服务器实现，服务器 200 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器（central processing units, CPU）22（例如，一个或一个以上处理器）和存储器 232，一个或一个以上存储应用程序 242 或数据 244 的存储介质 230（例如一个或一个以上海量存储设备）。其中，存储器 232 和存储介质 230 可以是短暂存储或持久存储。存储在存储介质 230 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对服务器中的一系列指令操作。更进一步地，中央处理器 22 可以设置为与存储介质
30 230 通信，在服务器 200 上执行存储介质 230 中的一系列指令操作。

35 服务器 200 还可以包括一个或一个以上电源 222，一个或一个以上有线或无线网络接口 250，一个或一个以上输入输出接口 258；或，一个或一个以上操作系统 241，例如 Windows

Server™, Mac OS X™, Unix™, Linux™, FreeBSD™ 等等。

本申请实施例中, 中央处理器 22, 用于执行本申请实施例中描述的任务求解方法。

应理解, 本申请实施例提供的任务求解方法也可以作为求解器部署在端侧的终端设备上, 这里并无限定。

5 本申请实施例中待求解模型可以用于解决调度问题。调度问题是大型制造/物流/生产等环
节中最常见的问题之一, 并且在不同的场景下, 调度总是有不同的意义。例如: 物流调
度主要是指在物流过程中, 物流公司根据待发货物的重量、去向、规格、加急程度等对所
属的车辆和人员进行合理的安排和调度。

10 而生产环境中的调度是根据不同产线中不同机器的产能以及生产需求, 在若干任务(job)
中完成对任务的排序以及任务和生产设备之间的匹配。即将多个任务分配至各条生产线中
的生产设备。

15 例如, 在作业车间调度(job-shop scheduling)的场景中, n 个工件在 m 台机器上加工,
每个工件有特定的加工工艺, 每个工件加工的顺序及每道工序所花时间给定, 安排工件在
每台机器上工件的加工顺序, 使得某种指标最优。这里不要求每个工件都在每个机器上执
行。

 例如, 在流水车间调度(flow-shop scheduling)的场景中, 该类调度问题要求每个任务
必须依次执行到每个阶段, 不涉及任务和阶段的匹配, 而主要是决定任务的执行顺序。防
止由于中间等待时间过长, 而造成整体的完成时间时长。

20 与一般的货物调度略有不同的是, 机场/大型制造工厂的工人/空姐排班(timetabling)
也是调度问题的一种, 这是由于这类问题的目标也是依照工人/空姐的工作特点以及场景需
要在不同的时间段内完成最优匹配。因此, 核心是排序以及最有分配, 而不局限“任务”
是人还是货物。一般来讲, 调度问题的目标是在给定任务数的前提下得到最小总工时
(makespan)所对应的排序。

25 同时, 调度问题在计算机中是分配工作所需资源的方法。资源可以指虚拟的计算资源,
如线程、进程或数据流; 也可以指硬件资源, 如处理器、网络连接或扩展卡。进行调度工
作的程序叫做调度器。调度器通常的实现使得所有计算资源都处于忙碌状态(在负载均衡
中), 允许多位用户有效地同时共享系统资源, 或达到指定的服务质量。

30 很多的调度问题(如排产、线体调度、加工网络布局等)都可以建模成一个数学问题
来求解, 线性规划(linear programming, LP)是其中使用最广的一类建模方法。目前, 线
性规划求解器底层依赖的算法中, 单纯型法是目前使用最为广泛的算法, 并且也是各个线
性规划求解器优化的比较多的一类算法。

35 对于单纯型算法, 通常, 算法先选择一个初始可行基解, 然后检查这个解是否已经达
到了最优解, 随后会对基矩阵做LU分解来计算基矩阵的逆; LU分解每K次迭代才会做一次,
其余的迭代都会用增量的方法来更新L矩阵和U矩阵的更新。然后, 算法会根据启发式的规
则选择入基变量和出基变量来更新基变量。不停执行这样的循环, 直到找到最优解, 或者
发现问题状态异常, 整个算法退出。

 从求解器的使用来分, 可以分成两种: 标准求解和连续求解。标准求解指求解器在没

有任何先验的情况下进行一次完整的求解，会分别经过预求解，求解，后处理等步骤，标准流程是在Simplex算法的前后分别增加了一个预求解模块和后处理模块。预求解模块会对模型进行化简，根据模型结构不同，化简的程度也不同，通常能将模型规模（变量数量、约束数量）减少30%~70%，这样Simplex可以求解一个相对简单的模型，从而加速求解。后处理模块把预求解生成的简单模型的解映射回原问题，最终得到原模型的解。

连续求解指前面已经对模型经过一次求解得到了最优解，随后对模型进行部分修改，比如增加一些变量约束或者修改已有变量的上下界，再次求解。这样基于上次最优解进行重求解的过程称为连续求解。

然而，上述描述的现有的标准求解流程没有利用任何先验知识来构造初始可行基，因此构造的初始可行基比较简单，从简单可行基出发进行单纯型算法迭代需要的迭代数过多，时间过长。且连续求解时候对于初始变量状态的调整方面做的过于直接，使得连续求解时找到初始的可行基解花费很长的时间。本申请实施例提供的任务求解方法可以基于先验知识来构造初始可行基，可以降低求解时间。

为了使得本申请更加的清楚，首先对本申请提到的部分概念和处理流程作简单介绍。

线性规划（Linear Programming, LP）：是运筹学中研究较早、发展较快、应用广泛、方法较成熟的一个重要分支，它是辅助人们进行科学管理的一种数学方法。研究线性约束条件下线性目标函数的极值问题的数学理论和方法。

约束：Constraints，是数学规划问题中的约束，即对决策变量的数值要求。

基解：Basic Solution，在约束方程组系数矩阵中找到一个基，令这个基的非基变量为零，再求解这个m元线性方程组就可得到唯一的解，这个解称之为线性规划的基本解。

基可行解：Basic Feasible Solution，基可行解即基本可行解的简称，是处理线性规划的基本概念。满足非负条件的基本解称为基可行解。

单纯形法：Simplex，单纯形法是求解线性规划问题最常用、最有效的算法之一。单纯形法的基本思路是：先找出可行域的一个顶点，根据一定规则判断其是否最优；若否，则转换到与之相邻的另一顶点，并使目标函数值更优；如此下去，直到找到某最优解为止。

下面结合本申请实施例中的附图对本申请实施例进行描述。本申请的实施方式部分使用的术语仅用于对本申请的具体实施例进行解释，而非旨在限定本申请。参见图3，图3为本申请实施例提供的一种任务求解方法，所述方法包括：

301、获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束。

在一种可能的实现中，步骤301的执行主体可以为服务器，例如终端设备可以将第一线性规划任务作为待求解模型传递至服务器，进而服务器可以获取第一线性规划任务。又例如，终端设备可以将第一线性规划任务作为求解待求解模型时所使用的的先验信息，也就是将包括第一线性规划任务在内的至少一个历史模型传递至服务器，进而服务器可以获取包括第一线性规划任务在内的至少一个历史模型。

通常，线性规划模型可以包括目标函数和约束条件，其中，目标函数是指根据待优化的目标和影响该目标的变量所设计的函数。例如，在排产问题中，整个排产的目标通常是

在满足所有资源约束的情况下，找出一个最好的加工计划，使得需求的满足率最高，同时整体的成本最小（例如成本可以包括但不限于加工成本、库存成本、转运成本），此时，该目标函数可以是用于表示满足率最大化以及成本最小化的函数。另外，约束条件是指在求解目标函数的过程中所要满足的其他限制条件。

- 5 在一种可能的实现中，所述第一线性规划任务用于为至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

例如，在产品生产的场景中，待调度任务可以是待生产的产品，在人员调度的场景中，待调度任务可以是待生产的人等等，本申请实施例并不限定。

- 10 在产品生产的场景中，多个可调度资源组中的每个可调度资源组可以为生产线，例如在手机的场景场景中，多个可调度资源组中的每个可调度资源组可以为一种手机组件的生产线，例如可以是电池的生产线、外壳的生产线、芯片的生产线等等，相应的，每个可调度资源组可以包括多个可调度资源，多个可调度资源中的每个可调度资源为所述生产线中的生产设备，例如，电池生产线可以包括多个电池生产设备，外壳生产线可以包括多个外壳生产设备，这里并不限定。
- 15

- 在人员调度的场景中，多个可调度资源组中的每个可调度资源组可以为时间段，例如在人员调度的场景中，多个可调度资源组中的每个可调度资源组可以为一天，例如可以是周一、周二、周三或者是一些月份的某一天等等，相应的，每个可调度资源组可以包括多个可调度资源，多个可调度资源中的每个可调度资源为时间段中的子时间段，例如，某一天可以包括多个小时、多个分钟或者其他多个子时间段，这里并不限定。
- 20

其中，第一规划约束可以包括求解变量，以及各个求解变量需要满足的约束。

302、获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度。

- 25 在一种可能的实现中，第一线性规划任务可以作为待求解模型。在求解第一线性规划任务时，可以首先获取各个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度。

- 通常情况下，线性规划任务中约束的数量较多时，求解线性规划任务所需的求解时间很长，因此可以首先选择线性规划任务中的部分约束，并基于选择的部分约束进行求解，并将求解结果（选择的部分约束中求解变量的状态）赋予线性规划任务，相当于将求解结果作为线性规划任务的初始值，并对线性规划任务进行求解，如果选择的部分约束的求解结果与求解线性规划任务后的求解结果基本一致（或者描述为较为接近，这里的求解结果基本一致可以理解为相同的约束中的求解变量的参数值基本一致），则求解线性规划任务时所需的迭代次数较少，也就是可以提高线性规划任务的求解速度。
- 30

- 35 因此，在求解第一线性规划任务之前，需要选择一部分约束进行求解，且得到的求解结果与第一线性规划任务的求解结果较为接近。

本申请实施例中，基于重要性来作为从多个第一规划约束中选择部分约束的依据，其

中重要性可以指示对于降低所述第一线性规划任务的求解时间的贡献程度。其中，所谓指示对于降低所述第一线性规划任务的求解时间的贡献程度，可以理解为，在构建待求解模型的子模型时（也就是从待求解模型中的约束中选择一部分来构建子模型），在将子模型求解的结果作为待求解模型的初始值的情况下，子模型包括该第一线性规划任务对于降低求解待求解模型的求解时间的贡献程度。

5 具体的，可以将重要性作为采样概率对多个第一规划约束进行采样可以得到一个子任务（第二规划任务），该子任务中的规划约束为对于降低所述第一线性规划任务的求解时间的贡献程度较高的约束，因此，对第二规划任务求解后得到的求解结果与对第二规划任务求解后得到的求解结果基本一致（或者描述为较为接近，这里的求解结果基本一致可以理解为相同的约束中的求解变量的参数值基本一致）。

10 在一种可能的实现中，该重要性可以基于先验信息得到，其中先验信息可以为多个和待求解模型（第一线性规划任务）同类的一个或多个模型，所谓同类，可以理解为线性规划任务包括的约束类型、数量一致或者基本一致，也就是说第一线性规划任务和先验信息中的一个或多个模型包括的约束类型、数量一致或者基本一致。

15 在确定各个约束的重要性时，可以直接对先验信息中各个线性规划任务进行求解，得到求解各个线性规划任务的求解时间（若线性规划任务已经被求解过，也可以直接获取到求解时间），然后对先验信息中的一个或多个线性规划模型进行子模型的提取、求解，并基于求解结果对线性规划模型进行初始化，然后对线性规划模型进行求解，以得到另一个求解时间，通过比对直接求解各个线性规划任务的求解时间，和对基于子模型求解结果进行
20 初始化后的线性规划任务的求解时间，就可以知晓子模型中的约束对于降低线性规划任务的求解时间的贡献程度。

其中，参阅图4，为了能够量化上述重要性，可以对待求解模型的各个约束赋予一个概率（该概率可以为上述重要性的量化），并基于概率对待求解模型进行采样，以构建子模型，并基于比对直接求解各个线性规划任务的求解时间，和对基于子模型求解结果进行初始化
25 后的线性规划任务的求解时间，来更新概率，迭代一定的次数，就可以得到用于准确描述约束对于降低线性规划任务的求解时间的重要性信息（关于重要性的更新方式将在候选实施例中描述）。

基于上述重要性采样得到的子模型（例如本申请实施例中的第二线性规划任务）中的各个规划约束对降低线性规划任务的求解时间的贡献程度很高，因此，基于第二线性规划任务的求解结果初始化后的第一线性规划模型的求解过程的时间可以被大大缩短。

30 应理解，本申请实施例中的第一线性规划模型可以为计算各个第一规划约束的重要性时采用的先验信息，在这种情况下，步骤302所获取的所述多个第一规划约束中每个第一规划约束的重要性可以为初始化的重要性，或者是迭代过程中的重要性（还未收敛到满足要求）。

35 应理解，计算各个第一规划约束的重要性时采用的先验信息可以包括多个模型，第一线性规划任务可以为多个模型中的一个，且第一线性规划任务可以是多个模型中随机选择的。

示例性的，可以获取 N 个历史上跟待求解模型同质的模型 P ，比如在过去不同时间点的

- 13 -

同一类优化问题。并且 N 个模型一共有 C 类约束，每一类约束用 c_i 表示；预先求解每个问题 $P_i, i = 1, 2, \dots, N$ ，得到它们直接求解的时间为 T_{P_i} ；设每类约束的重要性为 $\alpha_{c_i}, i = 1, 2, \dots, C$ ，并且它们的初始值为 $\alpha_{c_i} = 0.5$ （或者设定为其他初始值）；选择一个模型 P_i ，并根据 α 的概率采样出子模型 P_i^j （包含了约束 $c_{j_1}, \dots, c_{j_k}$ ），求解这个子模型，然后把子模型解作为原问题的

- 5 初始值求解原问题，记录下总的时间（包含求解子模型的时间） T_j ；更新 $\alpha_{c_{j_1}} \sim \alpha_{c_{j_k}}$ ，示例性的，可以基于如下公式更新： $\alpha_{c_i} = \alpha_{c_i} + \rho \times g(T_j, T_{P_i}, P_i^j, c_i)$ ；

其中 ρ 为步长， $g(\cdot)$ 为概率更新函数，重复上述步骤，直到重要性收敛。

在得到各个规划约束的重要性后，在对第一线性规划任务进行求解前，可以获取到第一线性规划任务中各个第一线性规划任务对应的重要性。

10

303、根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；

304、根据所述多个第一规划约束的子集构建第二线性规划任务；

- 15 在一种可能的实现中，重要性越高的第一规划约束，越大概率被采样作为第二线性规划任务的一部分，因此，第二线性规划任务可以包括所述多个第一规划约束中的部分第一规划约束。

- 20 在一种可能的实现中，第一线性规划模型为待求解模型或者用于更新重要性所使用的先验信息，可以将收敛的重要性作为各个约束的采样概率，并对多个第一规划约束进行采样，以获取第二线性规划任务，基于上述重要性采样得到的子模型（例如本申请实施例中的第二线性规划任务）中的各个规划约束对降低线性规划任务的求解时间的贡献程度很高，因此，基于第二线性规划任务的求解结果初始化后的第一线性规划模型的求解过程的时间可以被大大缩短。

305、对所述第二线性规划任务进行求解，得到第一求解结果。

- 25 在一种可能的实现中，所述第二线性规划任务包括 M 个求解变量，所述第一求解结果可以包括各个求解变量的参数值，其中这里的参数值可以为最优解的数值或者是求解变量的状态。在将所述第一求解结果作为所述第一线性规划任务的初始值时，具体可以将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的 M 个求解变量的参数值。也就是将第一求解结果中求解变量的参数值赋予第一线性规划任务，而第一线性规划任务中除了第二线性规划任务中包括的规划约束之外的第一规划约束的初始值可以由求解器指定，这里并不限定。

30

306、将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。

在一种可能的实现中, 求解器可以将第一求解结果读入第一线性规划任务, 随后第一线性规划任务中的求解变量 v_i 的初始状态赋为 s_i , 并对初始化后的所述第一线性规划任务进行求解, 以得到第二求解结果。

5 在一种可能的实现中, 第一线性规划模型可以为待求解模型, 此时第二求解结果可以为待求解模型的求解结果, 服务器在得到第二求解结果之后, 可以将第二求解结果回传至终端设备。

10 在一种可能的实现中, 所述得到第二求解结果之后, 可能还会新加约束(例如第二规划约束), 可以获取第二规划约束, 所述第二规划约束包括目标求解变量以及松弛变量, 将所述第二规划约束增加至所述第一线性规划任务, 以得到更新后的所述第一线性规划任务, 并求解所述更新后的第一线性规划任务。也就是可以给新添加的约束中涉及到的新变量根据其实际含义设置一个合理的bound(数值范围), 并修改这些变量的初始状态值, 将状态改到步骤一设置的bound上, 使得初始解对于新添加的约束是可行的, 进而加速线性规划的求解。

15 本申请可以在连续求解场景中针对新添加的松弛变量, 预先设置变量上界和初始状态, 从而加速连续求解。

在一种可能的实现中, 第一线性规划模型可以为用于更新重要性所使用的先验信息, 在这种情况下, 可以基于比对直接求解各个线性规划任务的求解时间, 和对基于子模型求解结果进行初始化后的线性规划任务的求解时间, 来更新概率, 迭代一定的次数, 就可以得到用于准确描述约束对于降低线性规划任务的求解时间的重要性信息。

20 具体的, 可以获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间, 所述第一求解时间和所述第三求解时间的加和可以作为上述基于子模型求解结果进行初始化后的线性规划任务的求解时间, 并对所述第一线性规划任务进行求解, 得到求解所述第一线性规划任务的第三求解时间, 其中, 第三求解时间可以认为是直接求解第一线性规划任务的求解时间, 根据所述第一求解时间和所述第三求解时间的加和, 相对于所述第一求解时间的减少程度, 就可以更新所述每个第一规划约束的重要性, 其中, 更新后的所述重要性与所述减少程度正相关。

25 应理解, 还可以将采样得到第二线性规划任务的过程所需的时间也作为基于子模型求解结果进行初始化后的线性规划任务的求解时间的一部分, 也就是可以获取对所述多个第一规划约束进行采样的采样时间, 并根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和, 相对于所述第一求解时间的减少程度, 更新所述每个第一规划约束的重要性。

30 例如, 构造子模型花费的时间记为 T_a 。对这个子模型(包含 $c_{j_1}, \dots, c_{j_K}$ 一共 K 类约束)调用线性规划求解器进行求解, 求解得到的子模型的最优解 S , 求解花费时间记为 T_b 。最后将子模型的解作为初始解带入原模型 P_i 中, 再求解一次 P_i , 这个步骤花费的时间为 T_c 。最后将这次重要性采样求解得到的总时间记为 $T_j = T_a + T_b + T_c$ 。

示例性的, 可以对于每一类约束 c_{j_1} , 可以执行下面的公式来进行更新: $\alpha_{c_{j_1}} =$

$\min(\alpha_{c_{j1}} + \rho \times \frac{T_{P_i} - T_j}{K}, 1)$ 。其中， ρ 可以为预先设置的超参。

上述基于根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性可以为对重要性的一次迭代过程，可以基于其他先验信息或者对第一线性规划模型多次执行上述迭代过程来更新重要性，以得到收敛的重要性。

服务器可以基于收敛后的重要性，对待求解模型进行求解。

例如，若待求解模型为目标线性规划任务，可以基于收敛后的重要性，对目标线性规划任务中的规划约束进行采样，以得到子模型，并对子模型进行求解，之后将求解结果作为目标线性规划任务的初始值，再对目标线性规划任务进行求解，以得到最终的求解结果，服务器可以将上述得到的最终求解结果反馈至终端设备。

以待求解模型为第三线性规划任务为例，可以获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；

获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；

根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；对所述第四线性规划任务进行求解，得到第三求解结果；将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三线性规划任务进行求解，得到第四求解结果。

示例性的，以多工厂排产问题为例，整个排产的目标通常是在满足所有资源约束的情况下，找出一个最好的加工计划，使得需求的满足率最高，同时整体的成本最小（加工成本、库存成本、转运成本）。图5展示了一个排产样例，假设客户下单了2000台PC，1000台主机，800台笔记本。然后有两家工厂可以加工这三类产品（每天最多加工1000件产品）。加工1台PC必须要先有一台主机，并且组装成PC只能在第一家工厂做。图中已经展示了一个最优的排产方案，即第一天在两家工厂各加工1000台主机，并且将第二家工厂的主机运往第一家工厂；第二天在第一家工厂把第一天加工完的主机组装成PC，然后再第二家工厂再加工1000台主机；第三天在第一家工厂使用第一天从第二家工厂运出的主机组装成1000台PC，最后在第二家工厂加工800台笔记本。这样，所有的需求都能满足，并且总的成本最低。多工厂排产本质是一个整数规划的问题，但是由于整数规划的求解复杂度过高，对于这样超大规模的模型，很难在规定的时间内求得最优解。所以通常的做法是将问题松弛成线性规划，然后再通过一些近似的方法，把线性规划求解得到的实数解调整成整数解。

首先收集历史三个月的排产模型，构成90个模型左右的训练数据集。同时对于排产问题主要库存约束、产能约束、配对约束、特控约束、SR约束、订单需求延期约束、预测需求延期约束、最大替代约束等18类约束。

对这90个模型先用线性规划求解器求解一次，记录求解的时间 $T_{P_i}, i = 1, \dots, 90$ ，初始化

约束的重要性概率 $\alpha_{c_i} = 0.5$ ，选择一个模型 P_i ，然后根据当前的 α 采样出一个子模型 P_i^j ，采样的方法为对于每类约束，随机生成一个随机数 r ，如果 $r \leq \alpha_{c_i}$ 。那么这类约束和它对应的变量都被纳入子模型中，构造子模型花费的时间记为 T_a 。对这个子模型（包含 $c_{j_1}, \dots, c_{j_K}$ 一共 K 类约束）调用线性规划求解器进行求解，求解得到的子模型的最优解 S ，求解花费时间记为 T_b 。最后将子模型的解作为初始解带入原模型 P_i 中，再求解一次 P_i ，这个步骤花费的时间为 T_c 。最后将这次重要性采样求解得到的总时间记为 $T_j = T_a + T_b + T_c$ 。中记录的 T_j 更新重要性概率 α 。对于子模型中的每一类约束 c_{j_1} ，执行下面的更新公式： $\alpha_{c_{j_1}} = \min(\alpha_{c_{j_1}} + \rho \times \frac{T_{P_i} - T_j}{K}, 1)$ ；继续选择模型进行更新操作，直到更新一定次数为止，得到训练好的 α ；拿到一个当前最新时刻的排产模型，根据历史数据中训练好的 α ，先采样一个子模型，随后求解这个子模型得到子模型的解 S ，把 S 作为初始值带入到原模型中，求解原模型得到这个排产问题的一个实数解。

针对排产问题中的最小批量约束，也就是每天的加工量要么等于0，要么大于等于某个常数 C 。由于建立线性规划模型时候无法考虑这个约束，因此这里可以使用连续求解的方法，对已经求解完的模型添加形如下面样式的约束： $x_i + s_j \geq C$ ；其中 x_i 是前面标准求解得到的某个大于0的变量， s_j 是新添加的松弛变量。同时将 s_j 加入目标函数中。此时再求解一次加了约束的模型，达到的效果是 x_i 能够尽量大于等于 C 。中新增的松弛变量 s_j ，设置变量的UpperBound $s_j^{UB} = C$ ，同时在求解前设置 $\text{status}(s_j) = \text{atUpperBound}$ 。求解修改过松弛变量初始状态的模型，得到解。经过的步骤，就可以完成排产问题模型的求解。

表1展示了本申请实施例的求解提升效果，相比直接求解，利用本申请实施例的求解方法效率提升平均有20%左右。

表1

20Days	Warm Start				Direct Solve		Improve Rate
	pick_sub	solve_sub	transfer_time	warmStart_total_time	direct_solve_total_time	percentage	
20200503	0.978747	262.022	53.472	345.541	421.706	0.179425951	
20200504	1.14935	290.433	166.303	433.265	560.373	0.128923059	
20200505	1.13718	253.403	65.392	344.15	427.113	0.197147195	
20200506	1.12473	258.013	122.363	440.941	569.26	0.540910822	
20200507	1.19738	259.822	65.832	344.743	431.39	0.177676812	
20200508	1.13381	375.293	158.473	393.484	1999	0.401516	
20200509	1.11446	269.312	95.242	369.103	612.242	0.370995456	
20200510	1.03653	562.692	97.532	392.113	537.186	0.269993069	
20200511	1.04711	269.552	47.632	319.194	435.323	0.22082224	
20200512	1.14782	314.22	105.692	435.993	463.637	0.01660597	
20200513	1.13379	309.352	133.812	433.44	1999	0.50456	
20200514	1.17397	364.712	100.512	495.501	762.722	0.348647345	

本申请实施例提供了一种任务求解方法，所述方法包括：获取第一线性规划任务，所

述第一线性规划任务包括多个第一规划约束；获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度；根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；根据所述多个第一规划约束的子集构建第二线性规划任务；将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。通过上述方式，一方面选取第一线性规划任务中的关键约束构建子模型（第二线性规划任务），然后利用子模型的解来第一线性规划任务从而加速求解过程。另一方面，在构建子模型时，基于重要性对第一线性规划任务进行采样，由于重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度，使采样得到的子模型的解接近于最优解，从而加速第一线性规划任务的求解过程。

本申请实施例还提供一个系统，其中，系统可以包括终端设备以及服务器，其中，终端设备可以发送待求解模型（第一线性规划任务）以及包括多个历史模型在内的先验信息至服务器，服务器可以基于包括多个历史模型在内的先验信息计算得到收敛的重要性，并根据重要性采样第一线性规划任务，并执行上述实施例中步骤301至步骤303的步骤，以得到第二求解结果，并将第二求解结果回传至终端设备。

此外终端设备可以发送待求解模型以及包括多个历史模型（包括第一线性规划任务）在内的先验信息至服务器，服务器可以基于包括多个历史模型在内的先验信息计算得到收敛的重要性（计算过程可以参照上述实施例中关于更新重要性相关的描述），并根据重要性采样待求解模型，以及求解待求解模型。

参照图6，图6为本申请实施例提供的一种任务求解装置的结构示意，如图6所示，所述装置600可以包括：

获取模块601，用于获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；

获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度；

其中，关于获取模块601的具体描述可以参照上述实施例中步骤301和302的描述，这里不再赘述。

采样模块602，用于根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；

根据所述多个第一规划约束的子集构建第二线性规划任务；

其中，关于采样模块602的具体描述可以参照上述实施例中步骤303和步骤304的描述，这里不再赘述。

求解模块603，用于对所述第二线性规划任务进行求解，得到第一求解结果；

将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一

线性规划任务进行求解，得到第二求解结果。

其中，关于求解模块603的具体描述可以参照上述实施例中步骤305和306的描述，这里不再赘述。

在一种可能的实现中，所述第一线性规划任务和所述第二线性规划任务均包括规划目标。

在一种可能的实现中，所述获取模块，还用于：

获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间；

所述求解模块，还用于对所述第一线性规划任务进行求解，得到求解所述第一线性规划任务的第三求解时间；

所述装置还包括：

重要性更新模块，用于根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，其中，更新后的所述重要性与所述减少程度正相关。

在一种可能的实现中，所述获取模块，还用于：

获取对所述多个第一规划约束进行采样的采样时间；

所述重要性更新模块，具体用于：

根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性。

在一种可能的实现中，所述获取模块，还用于：

获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；

获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；

所述采样模块，还用于：

根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；

所述求解模块，还用于：

对所述第四线性规划任务进行求解，得到第三求解结果；

将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三线性规划任务进行求解，得到第四求解结果。

在一种可能的实现中，所述第二线性规划任务包括M个求解变量，所述第一求解结果包括各个求解变量的参数值；

所述求解模块，具体用于：

将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的M个求解变量的参数值。

在一种可能的实现中，所述第一线性规划任务用于为至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

在一种可能的实现中，所述获取模块，还用于：

- 5 在所述得到第二求解结果之后，获取第二规划约束，所述第二规划约束包括目标求解变量以及松弛变量；

所述求解模块，还用于将所述第二规划约束增加至所述第一线性规划任务，以得到更新后的所述第一线性规划任务，并求解所述更新后的第一线性规划任务。

- 10 接下来介绍本申请实施例提供的一种终端设备，请参阅图 7，图 7 为本申请实施例提供的终端设备的一种结构示意图，终端设备 700 具体可以表现为手机、平板、笔记本电脑、智能穿戴设备、服务器等，此处不做限定。其中，终端设备 700 上可以部署有图 6 对应实施例中所述的任务求解装置，用于实现图 6 对应实施例中任务求解的功能。具体的，终端设备 700 包括：接收器 701、发射器 702、处理器 703 和存储器 704 (其中终端设备 700 中的处理器 703 的数量可以一个或多个，图 7 中以一个处理器为例)，其中，处理器 703 可以包括应用处理器 7031 和通信处理器 7032。在本申请的一些实施例中，接收器 701、发射器 702、处理器 703 和存储器 704 可通过总线或其它方式连接。

- 15 存储器 704 可以包括只读存储器和随机存取存储器，并向处理器 703 提供指令和数据。存储器 704 的一部分还可以包括非易失性随机存取存储器 (non-volatile random access memory, NVRAM)。存储器 704 存储有处理器和操作指令、可执行模块或者数据结构，或者它们的子集，或者它们的扩展集，其中，操作指令可包括各种操作指令，用于实现各种操作。

- 20 处理器 703 控制终端设备的操作。具体的应用中，终端设备的各个组件通过总线系统耦合在一起，其中总线系统除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都称为总线系统。

- 25 上述本申请实施例揭示的方法可以应用于处理器 703 中，或者由处理器 703 实现。处理器 703 可以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 703 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 703 可以是通用处理器、数字信号处理器 (digital signal processing, DSP)、微处理器或微控制器，还可进一步包括专用集成电路 (application specific integrated circuit, ASIC)、现场可编程门阵列 (field-programmable gate array, FPGA) 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。该处理器 703 可以实现或者执行本申请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该处理器也可以是任何常规的处理器等。结合本申请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。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 704，处理器 703 读取存储器 704 中的
- 30
- 35

信息，结合其硬件完成上述方法中关于终端设备的步骤。

接收器 701 可用于接收输入的数字或字符信息，以及产生与终端设备的相关设置以及功能控制有关的信号输入。发射器 702 可用于通过第一接口输出数字或字符信息；发射器 702 还可用于通过第一接口向磁盘组发送指令，以修改磁盘组中的数据；发射器 702 还可以包括显示屏等显示设备。

本申请实施例中，在一种情况下，处理器 703，用于执行上述实施例中的终端设备执行的步骤。

本申请实施例还提供了一种服务器，请参阅图 8，图 8 是本申请实施例提供的服务器一种结构示意图，具体的，服务器 800 由一个或多个服务器实现，服务器 800 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器（central processing units, CPU）88（例如，一个或一个以上处理器）和存储器 832，一个或一个以上存储应用程序 842 或数据 844 的存储介质 830（例如一个或一个以上海量存储设备）。其中，存储器 832 和存储介质 830 可以是短暂存储或持久存储。存储在存储介质 830 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对服务器中的一系列指令操作。更进一步地，中央处理器 88 可以设置为与存储介质 830 通信，在服务器 800 上执行存储介质 830 中的一系列指令操作。

服务器 800 还可以包括一个或一个以上电源 826，一个或一个以上有线或无线网络接口 850，一个或一个以上输入输出接口 858；或，一个或一个以上操作系统 841，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 等等。

本申请实施例中，中央处理器 88，用于执行上述实施例任务求解方法相关的步骤。

本申请实施例中还提供一种包括计算机程序产品，当其在计算机上运行时，使得计算机执行如前述终端设备所执行的步骤，或者，使得计算机执行如前述服务器所执行的步骤。

本申请实施例中还提供一种计算机可读存储介质，该计算机可读存储介质中存储有用于进行信号处理的程序，当其在计算机上运行时，使得计算机执行如前述终端设备所执行的步骤，或者，使得计算机执行如前述服务器所执行的步骤。

本申请实施例提供的终端设备、服务器或终端设备具体可以为芯片，芯片包括：处理单元和通信单元，所述处理单元例如可以是处理器，所述通信单元例如可以是输入/输出接口、管脚或电路等。该处理单元可执行存储单元存储的计算机执行指令，以使终端设备内的芯片执行上述实施例描述的数据处理方法，或者，以使服务器内的芯片执行上述实施例描述的数据处理方法。可选地，所述存储单元为所述芯片内的存储单元，如寄存器、缓存等，所述存储单元还可以是所述无线接入设备端内的位于所述芯片外部的存储单元，如只读存储器(read-only memory, ROM)或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器(random access memory, RAM)等。

具体的，请参阅图 9，图 9 为本申请实施例提供的芯片的一种结构示意图，所述芯片可以表现为神经网络处理器 NPU 900，NPU 900 作为协处理器挂载到主 CPU (Host CPU) 上，由 Host CPU 分配任务。NPU 的核心部分为运算电路 903，通过控制器 904 控制运算电路 903 提取存储器中的矩阵数据并进行乘法运算。

在一些实现中，运算电路 903 内部包括多个处理单元 (Process Engine, PE)。在一些实现中，运算电路 903 是二维脉动阵列。运算电路 903 还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数学运算的其它电子线路。在一些实现中，运算电路 903 是通用的矩阵处理器。

5 举例来说，假设有输入矩阵 A，权重矩阵 B，输出矩阵 C。运算电路从权重存储器 902 中取矩阵 B 相应的数据，并缓存在运算电路中每一个 PE 上。运算电路从输入存储器 901 中取矩阵 A 数据与矩阵 B 进行矩阵运算，得到的矩阵的部分结果或最终结果，保存在累加器 (accumulator) 908 中。

10 统一存储器 906 用于存放输入数据以及输出数据。权重数据直接通过存储单元访问控制器 (Direct Memory Access Controller, DMAC) 905，DMAC 被搬运到权重存储器 902 中。输入数据也通过 DMAC 被搬运到统一存储器 906 中。

BIU 为 Bus Interface Unit 即，总线接口单元 910，用于 AXI 总线与 DMAC 和取指存储器 (Instruction Fetch Buffer, IFB) 909 的交互。

15 总线接口单元 910 (Bus Interface Unit, 简称 BIU)，用于取指存储器 909 从外部存储器获取指令，还用于存储单元访问控制器 905 从外部存储器获取输入矩阵 A 或者权重矩阵 B 的原数据。

DMAC 主要用于将外部存储器 DDR 中的输入数据搬运到统一存储器 906 或将权重数据搬运到权重存储器 902 中或将输入数据数据搬运到输入存储器 901 中。

20 向量计算单元 907 包括多个运算处理单元，在需要的情况下，对运算电路 903 的输出做进一步处理，如向量乘，向量加，指数运算，对数运算，大小比较等等。主要用于神经网络中非卷积/全连接层网络计算，如 Batch Normalization(批归一化)，像素级求和，对特征平面进行上采样等。

25 在一些实现中，向量计算单元 907 能将经处理的输出的向量存储到统一存储器 906。例如，向量计算单元 907 可以将线性函数；或，非线性函数应用到运算电路 903 的输出，例如对卷积层提取的特征平面进行线性插值，再例如累加值的向量，用以生成激活值。在一些实现中，向量计算单元 907 生成归一化的值、像素级求和的值，或二者均有。在一些实现中，处理过的输出的向量能够用作到运算电路 903 的激活输入，例如用于在神经网络中的后续层中的使用。

30 控制器 904 连接的取指存储器 (instruction fetch buffer) 909，用于存储控制器 904 使用的指令；

统一存储器 906，输入存储器 901，权重存储器 902 以及取指存储器 909 均为 On-Chip 存储器。外部存储器私有于该 NPU 硬件架构。

其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，ASIC，或一个或多个用于控制上述程序执行的集成电路。

35 另外需说明的是，以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际

的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外，本申请提供的装置实施例附图中，模块之间的连接关系表示它们之间具有通信连接，具体可以实现为一条或多条通信总线或信号线。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CPU、专用存储器、专用元器件等来实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在可读取的存储介质中，如计算机的软盘、U 盘、移动硬盘、ROM、RAM、磁碟或者光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等等）执行本申请各个实施例所述的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。

所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘（Solid State Disk，SSD））等。

权 利 要 求

1. 一种任务求解方法，其特征在于，所述方法包括：
- 获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；
- 获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划
- 5 约束对于降低所述第一线性规划任务的求解时间的贡献程度；
- 根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；
- 根据所述多个第一规划约束的子集构建第二线性规划任务；
- 对所述第二线性规划任务进行求解，得到第一求解结果；
- 10 将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。
2. 根据权利要求1所述的方法，其特征在于，所述第一线性规划任务和所述第二线性规划任务均包括规划目标。
- 15 3. 根据权利要求1或2所述的方法，其特征在于，所述方法还包括：
- 获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间；
- 对所述第一线性规划任务进行求解，得到求解所述第一线性规划任务的第三求解时间；
- 20 根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，其中，更新后的所述重要性与所述减少程度正相关。
4. 根据权利要求3所述的方法，其特征在于，所述方法还包括：获取对所述多个第一规划约束进行采样的采样时间；
- 25 所述根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，包括：
- 根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，以得到更新后的每个第一
- 30 规划约束的重要性。
5. 根据权利要求3或4所述的方法，其特征在于，所述方法还包括：
- 获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；
- 35 获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；
- 根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第

—24—

四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；

对所述第四线性规划任务进行求解，得到第三求解结果；

将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三

5 线性规划任务进行求解，得到第四求解结果。

6. 根据权利要求1至5任一所述的方法，其特征在于，所述第二线性规划任务包括M个求解变量，所述第一求解结果包括各个求解变量的参数值；

所述将所述第一求解结果作为所述第一线性规划任务的初始值，包括：

10 将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的M个求解变量的参数值。

7. 根据权利要求1至6任一所述的方法，其特征在于，所述第一线性规划任务用于为至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

15

8. 根据权利要求1至7任一所述的方法，其特征在于，所述得到第二求解结果之后，所述方法还包括：

获取第二规划约束，所述第二规划约束包括松弛变量以及所述松弛变量的上界；

20 将所述第二规划约束增加至所述第一线性规划任务，以得到更新后的所述第一线性规划任务，并求解所述更新后的所述第一线性规划任务。

9. 一种任务求解装置，其特征在于，所述装置包括：

获取模块，用于获取第一线性规划任务，所述第一线性规划任务包括多个第一规划约束；

25

获取所述多个第一规划约束中每个第一规划约束的重要性，所述重要性表示第一规划约束对于降低所述第一线性规划任务的求解时间的贡献程度；

采样模块，用于根据所述重要性，对所述多个第一规划约束进行采样，以得到所述获取所述多个第一规划约束的子集，其中，所述重要性用于确定第一规划约束的采样概率；

30

根据所述多个第一规划约束的子集构建第二线性规划任务；

求解模块，用于对所述第二线性规划任务进行求解，得到第一求解结果；

将所述第一求解结果作为所述第一线性规划任务的初始值，并对初始化后的所述第一线性规划任务进行求解，得到第二求解结果。

35 10. 根据权利要求9所述的装置，其特征在于，所述第一线性规划任务和所述第二线性规划任务均包括规划目标。

— 25 —

11. 根据权利要求9或10所述的装置，其特征在于，所述获取模块，还用于：

获取对所述第二线性规划任务进行求解的第一求解时间、以及对初始化后的所述第一线性规划任务进行求解的第二求解时间；

所述求解模块，还用于对所述第一线性规划任务进行求解，得到求解所述第一线性规划任务的第三求解时间；

所述装置还包括：

重要性更新模块，用于根据所述第一求解时间和所述第三求解时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性，其中，更新后的所述重要性与所述减少程度正相关。

12. 根据权利要求11所述的装置，其特征在于，所述获取模块，还用于：

获取对所述多个第一规划约束进行采样的采样时间；

所述重要性更新模块，具体用于：

根据所述第一求解时间、所述第三求解时间以及所述采样时间的加和，相对于所述第一求解时间的减少程度，更新所述每个第一规划约束的重要性。

13. 根据权利要求11或12所述的装置，其特征在于，所述获取模块，还用于：

获取第三线性规划任务，所述第三线性规划任务包括所述多个第二规划约束，所述多个第二规划约束和所述多个第一规划约束的约束类型相同；

获取所述多个第二规划约束中每个第二规划约束的重要性，其中更新后的每个第一规划约束的重要性用于作为约束类型相同的第二规划约束的重要性；

所述采样模块，还用于：

根据所述每个第二规划约束的重要性，对所述多个第二规划约束进行采样，以获取第四线性规划任务，其中，所述每个第二规划约束的重要性用于确定第二规划约束的采样概率，所述第四线性规划任务包括所述多个第二规划约束中的部分第二规划约束；

所述求解模块，还用于：

对所述第四线性规划任务进行求解，得到第三求解结果；

将所述第三求解结果作为所述第三线性规划任务的初始值，并对初始化后的所述第三线性规划任务进行求解，得到第四求解结果。

14. 根据权利要求9至13任一所述的装置，其特征在于，所述第二线性规划任务包括M个求解变量，所述第一求解结果包括各个求解变量的参数值；

所述求解模块，具体用于：

将所述第一求解结果中各个求解变量的参数值作为所述第一线性规划任务中的M个求解变量的参数值。

15. 根据权利要求9至14任一所述的装置，其特征在于，所述第一线性规划任务用于为

至少一个待调度任务分配调度资源，所述第一规划约束为调度资源满足的约束，所述调度资源为生产线、生产设备或生产厂家。

16. 根据权利要求9至15任一所述的装置，其特征在于，所述获取模块，还用于：

5 在所述得到第二求解结果之后，获取第二规划约束，所述第二规划约束包括松弛变量以及所述松弛变量的上界；

所述求解模块，还用于将所述第二规划约束增加至所述第一线性规划任务，以得到更新后的所述第一线性规划任务，并求解所述更新后的第一线性规划任务。

10 17. 一种计算机存储介质，其特征在于，所述计算机存储介质存储有一个或多个指令，所述指令在由一个或多个计算机执行时使得所述一个或多个计算机执行权利要求1-8中任一项所述方法的操作。

15 18. 一种计算机程序产品，其特征在于，包括计算机可读指令，当所述计算机可读指令在计算机设备上运行时，使得所述计算机设备执行如权利要求1至8任一所述的方法。

19. 一种系统，包括至少一个处理器，至少一个存储器以及至少一个通信接口；所述处理器、所述存储器和所述通信接口通过通信总线连接并完成相互间的通信；

所述至少一个存储器用于存储代码；

20 所述至少一个处理器用于执行所述代码，以执行如权利要求1-8任一所述的任务求解方法，以得到求解结果；

所述至少一个通信接口，用于与设备或通信网络通信，以将所述求解结果发送至所述设备或通信网络。

25