

(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
2020 年 6 月 18 日 (18.06.2020)



(10) 国际公布号
WO 2020/119648 A1

(51) 国际专利分类号:
H04L 29/08 (2006.01)

(21) 国际申请号: PCT/CN2019/124084

(22) 国际申请日: 2019 年 12 月 9 日 (09.12.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811537034.9 2018 年 12 月 14 日 (14.12.2018) CN

(71) 申请人: 深圳先进技术研究院 (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) [CN/CN]; 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 杜明哲 (DU, Ming Zhe); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。王洋 (WANG, Yang); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。须成忠 (XU, Cheng

Zhong); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(74) 代理人: 北京市诚辉律师事务所 (BEIJING CHENGHUI LAW FIRM); 中国北京市朝阳区朝阳北路 99 号楼 2 单元 905, Beijing 100123 (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, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

(54) Title: COMPUTING TASK UNLOADING ALGORITHM BASED ON COST OPTIMIZATION

(54) 发明名称: 一种基于代价优化的计算任务卸载算法

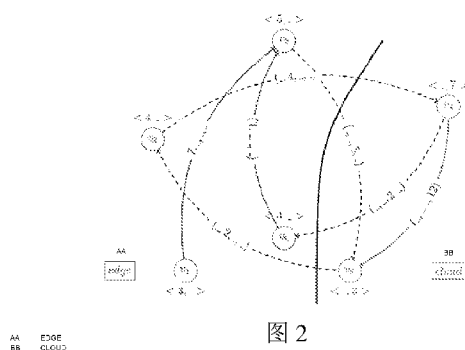


图 2

(57) Abstract: The present invention relates to the technical field of information, and in particular to a computing task unloading algorithm based on cost optimization. The algorithm mainly comprises the following steps: 1) constructing a new edge and cloud computing model, wherein the new edge and cloud computing model comprises three important computing costs: an execution cost of a computing task, a communication cost between computing tasks at the same end, and an asymmetric communication cost between computing tasks across ends; 2) extending the new edge and cloud computing model; 3) merging the computing costs; and 4) solving an optimized unloading strategy based on a greedy criterion. The algorithm of the present invention solves the problem of unloading optimization of a computing task in a combined frame of edge computing and cloud computing.

(57) 摘要: 本发明涉及信息技术领域, 具体而言, 涉及一种基于代价优化的计算任务卸载算法, 主要包括以下步骤: 1) 构建边云计算新模型; 所述边云计算新模型包括计算三种重要的代价: 计算任务的执行代价、同端计算任务之间的通信代价、跨端计算任务之间的非对称通信代价; 2) 对边云计算新模型进行扩展; 3) 合并计算代价; 4) 基于贪婪准则来求解优化卸载策略。本发明的算法解决了边缘计算和云计算的结合框架中, 计算任务的卸载优化问题。

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))。

一种基于代价优化的计算任务卸载算法

技术领域

本发明涉及信息技术领域，具体而言，涉及一种基于代价优化的计算任务卸载算法。

背景技术

在边云计算，也即边缘计算以及云计算的结合框架中，如何将任务从云端卸载到边缘端是提升计算任务处理效率、发挥边云计算框架的关键。边缘设备如移动基站、手机、网关等，靠近用户，也即靠近数据源，数据传输路径与传输时间相比于传统云计算而言大大缩短。但相比之下，边缘设备计算能力远低于云计算中心的计算能力，因此不能仅仅根据传输距离来确定任务在哪里执行。对于一批任务而言，如果将所有的数据都卸载到边缘设备上处理，其处理能力带来的问题比全部放在云端处理更严重；也不能随机将任务从云计算中心卸载边缘设备上处理，因为不同任务有不同的资源需求。比如有的任务需要大量计算资源，那么卸载到边缘设备上处理的性能就逊色于在云端处理的性能，因为数据中心的处理能力优势可以抵消掉数据传输带来的劣势；而有的任务不需要太多计算资源，但是对实时性有很高的要求，那么就适合卸载到边缘设备上处理。只有合理地将计算任务从云端卸载到边缘设备上，才能达到更好地结合边缘计算模型和云计算模型，既发挥边缘设备的优势，又降低云数据中心和网络带宽的压力，达到优化资源，提升效率的目的。

现有的边云计算模型一般采用简单的拓扑图形式表示一批计算任务，图中的点代表单个计算任务，任务一般有两个权重，分别代表计算任务在不同端的计算代价；而点之间的边代表两个计算任务之间存在通信，边上的权重代表了当两个任务跨端通信时的计算代价。任务卸载算法 MOCOP 将这种任务图作为输入参数，提出了基于最大流最小割定理的任务卸载算法。该算法将任务图分割成为两个部分，一部分在边缘执行，一部分在云端执行。最后的任务总执行代价为在边缘的任务的执行代价和在云端的任务的执行代价，以及边缘和云端

之间的通信代价的总和。基于上述的简单计算模型该算法给出了最优卸载策略，但是由于该算法并没有考虑到实际情况中边云模型同端通信代价以及跨端通信非对称等情况，因此该算法难以应用在实际的边云模型中，给出的卸载策略也并非最优卸载策略。

目前的计算任务卸载算法研究均基于一个简单的边云计算模型，该简化模型中没有考虑当任务同处于边缘设备或者云设备时的通信代价，而只考虑处于不同端的任务之间的通信代价，也即，当两个任务均位于云端或者均位于边缘端的时候，忽略计算任务之间的通信代价。而当一个计算任务在云端，一个计算任务在边缘端的时候，才考虑两个任务之间存在通信代价。此外，目前的计算任务卸载算法还基于一个简单的假设，也即，跨端通信代价是对称的：任务从云端到边缘端的通信与任务从边缘端到云端的通信代价是相等的。从而提出的任务卸载算法，也只优化任务计算代价以及对称的跨端通信代价。现实情况中，一个服务由多个应用组成，而多个应用分布在不同的机器上，当两个应用进行通信的时候，无论应用是位于同端的机器还是位于不同端的机器上，通信代价都是不可忽略的。此外，由于网络上下行带宽不一致，网络拥塞等因素的存在，跨端通信的代价也往往非对称。因此当前的计算任务卸载算法在实际的边云计算模型中无法给出精确的卸载策略。

发明内容

为解决上述背景技术中存在的问题，本发明提出一种基于代价优化的计算任务卸载算法，其更适用于实际情况中的在边和云平台之间的任务卸载，给出的卸载策略也能够大幅度优化任务的整体代价，也即任务的执行代价、通信代价和卸载代价。

本发明解决上述问题的技术方案是：一种基于代价优化的计算任务卸载算法，其特殊之处在于，包括以下步骤：

1) 构建边云计算新模型

所述边云计算新模型包括计算三种重要的代价：

- 1.1) 计算任务的执行代价；
- 1.2) 同端计算任务之间的通信代价；
- 1.3) 跨端计算任务之间的非对称通信代价；
- 2) 对边云计算新模型进行扩展；
- 3) 合并计算代价；
- 4) 基于贪婪准则来求解优化卸载策略。

进一步地，上述步骤 1.1) 计算任务的执行代价，具体为：

给定任务图 G ，将任务的执行代价 $C_{comp}(G)$ 表达为如下形式：

$$C_{comp}(G) = \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \quad (5.1),$$

其中 w_i^E 和 w_i^C 分别表示任务 i 在边缘 (E:edge) 和在云端 (C:cloud) 的执行代价，布尔变量 $x_i = 0$ 的时候，代表任务 i 保留在云端执行， $x_i = 1$ 的时候，代表任务 i 卸载到边缘执行。

进一步地，上述步骤 1.2) 及 1.3) 计算任务之间的通信代价，具体为：

用 $\langle w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4 \rangle$ 表示同端通信代价和跨端不对称的通信代价，也即任务 i 到任务 j 从边缘到边缘、边缘到云端、云端到边缘、边缘到边缘这四种通信代价；这四种通信代价统一用 $C_{comm}(G)$ 表示为如下形式：

$$C_{comm}(G) = \sum_{i,j} (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) \quad (5.2)。$$

进一步地，上述步骤 1.3) 计算任务的卸载代价，具体为：

任务 i 从云端卸载到边缘必须要考虑卸载代价，任务卸载代价 $C_{tran}(G)$ 记为：

$$C_{tran}(G) = \sum_{i=1}^n m_i x_i \quad (5.3),$$

为了进一步简化，将任务 i 的计算代价和卸载代价合并到一起：

$$\begin{aligned}
C_{comp}(G) &= \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i + m_i x_i) \\
&= \sum_{i=1}^n ((w_i^E + m_i) x_i + w_i^C \bar{x}_i) \\
&= \sum_{i=1}^n ((w_i^E)' x_i + w_i^C \bar{x}_i)
\end{aligned} \quad (5.4),$$

因此，原本总执行代价是计算代价、通信代价和卸载代价的总和，因为合并了计算代价和卸载代价，边缘计算中求解最优任务分配卸载可以表示为如下形式：

$$C_{total}(G) = \min_{x_i \in \{0,1\}} \{C_{comp}(G) + C_{comm}(G)\} \quad (5.5),$$

具体可表达为：

$$\begin{aligned}
&\min \left\{ \sum_{i,j}^n (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) + \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \right\} \\
&\quad s.t.: \\
&\quad x_i, x_j \in \{0,1\} \quad \forall i, j = 1, \dots, n.
\end{aligned} \quad (5.6)。$$

边云计算新模型中，代价优化的计算任务卸载问题最终表达为如上最小化问题。

进一步地，上述步骤2)中，对边云计算新模型进行扩展具体为：

边云计算新模型中，每两个任务节点 v_i, v_j 之间的代表通信的边 $e(v_i, v_j)$ 上均有四种通信代价可能 $w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4$ ，将四个权重的单条边扩展为四条边 $e(v_i^E, v_j^E)$ 、 $e(v_i^E, v_j^C)$ 、 $e(v_i^C, v_j^E)$ 和 $e(v_i^C, v_j^C)$ ，每条边上一个确定的通信代价。

进一步地，上述步骤3)中，合并计算代价具体为：

将节点上的两个计算代价合并到通信代价中：

首先，定义计算节点 v_i 的邻居节点（与 v_i 存在通信的任务节点）集合为 N_i ：

$$N_i = \{v_j | e(v_i, v_j) = 1\} \quad (5.7),$$

对于每一条边 $e(v_i^x, v_j^y)$ ($x, y \in \{E(\text{Edge}), C(\text{Cloud})\}$)，将其边权 $w(e(v_i^x, v_j^y))$ 改为：

$$w'(e(v_i^x, v_j^y)) = w(e(v_i^x, v_j^y)) + \frac{w_i^x}{|N_i|} + \frac{w_j^y}{|N_j|} \quad (5.8)。$$

进一步地，上述步骤4)中，基于贪婪准则来求解优化卸载策略，具体为：

定义 E_a 为所有边 $e(v_i^x, v_j^y) (x, y \in \{E(\text{Edge}), C(\text{Cloud})\})$ 的集合，每次从该集合中按照如下贪婪规则选取一条边：

$$\min \frac{w'(e(v_i^x, v_j^y))}{\sum_{e(v_i^x, v_j^y) \in E_a} w'(e(v_i^x, v_j^y))} \quad (5.9),$$

选取这条边后，将该边从 E_a 中删除，并将与边 $e(v_i^x, v_j^y)$ 不兼容的边，也即 $\bar{e}(v_i^x, v_j^y)$ ，从 E_a 中删除；

接下来循环该步骤，直至 E_a 为空。

进一步地，上述步骤4)中，

与 $e(v_i^x, v_j^y)$ 不兼容的边是指，比如：当任务 v_i^x 选择在边缘执行， v_j^y 在云端执行，那么与边 $e(v_i^E, v_j^C)$ 不兼容的边就是 $e(v_i^C, v_j^C)$ ，该边代表任务 v_i^x, v_j^y 均选择在云端执行；

$e(v_i^E, v_j^E)$ ，该边代表任务 v_i^x, v_j^y 均选择在边缘执行；以及 $e(v_i^C, v_j^E)$ ，该边代表任务 v_i^x 选择在云端执行， v_j^y 在边缘执行。

本发明的优点：

本发明是一种基于代价优化的计算任务卸载算法，解决了边云计算新模型中的计算任务卸载的优化问题，算法中加入了已有研究中忽略的同端通信代价，以及跨端通信非对称的特点，同时考虑了网络资源的异构性和同构性；本发明算法基于新模型而提出，更加适用于边云计算环境中的计算任务卸载；本发明提出了基于贪婪原则的任务卸载算法：当网络资源异构时，本发明利用贪婪算法求解近似最优卸载策略；该算法时间复杂度以及空间复杂度都较低，可以快速高效求解近似最优计算任务卸载策略；相较于现有技术，该算法可以同时优化计算代价以及同端、跨端的通信代价，以及任务卸载代价。

附图说明

图1是本发明实施例中的边云计算新模型中的任务例图，

图2是本发明实施例中的计算任务的卸载结果例图；

图 3 是本发明实施例中的一个任务例图；

图 4 是本发明算法步骤 2) 和步骤 3) 的图解；

图 5 是本发明算法进行步骤 4) 的图解一；

图 6 是本发明算法进行步骤 4) 的图解二；

图 7 是本发明算法进行步骤 4) 的图解三；

图 8 是本发明算法进行步骤 4) 的图解四。

具体实施方式

为使本发明实施方式的目的、技术方案和优点更加清楚，下面将结合本发明实施方式中的附图，对本发明实施方式中的技术方案进行清楚、完整地描述，显然，所描述的实施方式是本发明一部分实施方式，而不是全部的实施方式。基于本发明中的实施方式，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施方式，都属于本发明保护的范围。因此，以下对在附图中提供的本发明的实施方式的详细描述并非旨在限制要求保护的本发明的范围，而是仅仅表示本发明的选定实施方式。基于本发明中的实施方式，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施方式，都属于本发明保护的范围。

以下结合图 1、2 对本发明所提供的基于代价优化的计算任务卸载算法做详细说明，其中，图 1 是本发明实施例中的边云计算新模型中的任务例图，图 2 本发明实施例中的计算任务的卸载结果例图。由图 1 可知，边云计算新模型相较于简单的计算模型，在计算任务卸载策略时需要考虑更多的影响因素。

详细地，本发明基于代价优化的计算任务卸载算法，包括以下步骤：

1) 构建边云计算新模型，所述边云计算新模型包括计算三种重要的代价，即首先将代价优化的计算任务卸载问题表达为如下形式：

1.1) 计算任务的执行代价。

本实施例中，具体为：

给定任务图 G ，将任务的执行代价 $C_{comp}(G)$ 表达为如下形式：

$$C_{comp}(G) = \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \quad (5.1),$$

其中 w_i^E 和 w_i^C 分别表示任务 i 在边缘(E:edge)和在云端(C:cloud)的执行代价, 布尔变量 $x_i = 0$ 的时候, 代表任务 i 保留在云端执行, $x_i = 1$ 的时候, 代表任务 i 卸载到边缘执行。

1.2) 计算任务之间的通信代价。

本实施例中, 具体为:

具体为:

用 $\langle w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4 \rangle$ 表示同端通信代价和跨端不对称的通信代价, 也即任务 i 到任务 j 从边缘到边缘、边缘到云端、云端到边缘、边缘到边缘这四种通信代价; 这四种通信代价统一用 $C_{comm}(G)$ 表示为如下形式:

$$C_{comm}(G) = \sum_{i,j} (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) \quad (5.2)。$$

1.3) 计算任务之间的通信代价。

本实施例中, 具体为: 具体为:

任务 i 从云端卸载到边缘必须要考虑卸载代价, 任务卸载代价 $C_{tran}(G)$ 记为:

$$C_{tran}(G) = \sum_{i=1}^n m_i x_i \quad (5.3),$$

为了进一步简化, 将任务 i 的计算代价和卸载代价合并到一起:

$$\begin{aligned} C_{comp}(G) &= \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i + m_i x_i) \\ &= \sum_{i=1}^n ((w_i^E + m_i) x_i + w_i^C \bar{x}_i) \\ &= \sum_{i=1}^n ((w_i^E)' x_i + w_i^C \bar{x}_i) \end{aligned} \quad (5.4),$$

因此, 原本总执行代价是计算代价、通信代价和卸载代价的总和, 因为合并了计算代价和卸载代价, 边缘计算中求解最优任务分配卸载可以表示为如下形式:

$$C_{total}(G) = \min_{x_i \in \{0,1\}} \{C_{comp}(G) + C_{comm}(G)\} \quad (5.5),$$

具体可表达为:

$$\begin{aligned} \min \{ & \sum_{i,j}^n (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) + \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \} \\ \text{s.t.:} \\ & x_i, x_j \in \{0, 1\} \quad \forall i, j = 1, \dots, n. \end{aligned} \quad (5.6)。$$

以上可以看出,相比于简单的计算模型,边云计算新模型复杂度大大增加,不利于算法求解。故本发明继续对提出的算法进行简化、求解,包括三部分:

第一、二部分为简化计算模型。由于加入上述参数之后,任务图上代表计算任务的点具有两个点权,而代表计算任务之间通信的边存在四个边权(分别代表任务均在云设备上通信的代价,均在边缘设备上通信的代价,一个任务在云设备,一个任务在边缘设备上通信的代价,以及一个任务在边缘设备,一个任务在云设备上通信的代价),任务的两个计算代价依据任务的执行位置而定,任务之间的通信代价又依据两个任务的相对执行位置而定。由于情况多变复杂,难以在低时间空间复杂度的情况下进行处理,因此本算法很重要的一步就是要对计算模型进行简化,同时要保证各个参数不变。为了简化计算模型,本算法主要将任务之间的单个通信边扩展成为四条边,每条边代表一种通信可能性,同时每条边都有一个确定的边权,也即通信代价。此外,还将点权,也即任务的计算代价合并到边权中,如此一来将复杂的计算模型转换成了一个传统的简单的计算模型。

第三部分为求解计算任务卸载策略;本发明为了降低时间复杂度和空间复杂度,采用贪婪原则来实现任务卸载策略。

这里结合图 3-图 8 做出具体说明:

2) 扩展计算模型:

如图 3 所示,边云计算新模型中,每两个任务节点 v_i, v_j 之间的代表通信的边 $e(v_i, v_j)$ 上均有四种通信代价可能 $w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4$ (为了简便,这里图上并没有给出具体的四个值,而是用 $w(e(v_i, v_j))$ 来代替四种通信代价);将图 3 中具有四个权重可能的单条边,扩展为图 4 所示的四条边,也即 $e(v_i^E, v_j^E)$ 、 $e(v_i^E, v_j^C)$ 、

$e(v_i^C, v_j^E)$ 和 $e(v_i^C, v_j^C)$), 每条边代表一种通信可能, 以及一个确定的通信代价;

3) 合并计算代价:

由于每个计算节点都有两种计算代价, 本算法进一步将节点上的两个计算代价合并到通信代价中:

首先, 定义计算节点 v_i 的邻居节点 (与 v_i 存在通信的任务节点) 集合为 N_i :

$$N_i = \{v_j | e(v_i, v_j) = 1\} \quad (5.7),$$

对于每一条边 $e(v_i^x, v_j^y)$ ($x, y \in \{E(\text{Edge}), C(\text{Cloud})\}$), 将其边权 $w(e(v_i^x, v_j^y))$ 改为:

$$w'(e(v_i^x, v_j^y)) = w(e(v_i^x, v_j^y)) + \frac{w_i^x}{|N_i|} + \frac{w_j^y}{|N_j|} \quad (5.8);$$

4) 定义 E_a 为所有边 $e(v_i^x, v_j^y)$ ($x, y \in \{E(\text{Edge}), C(\text{Cloud})\}$) 的集合, 每次从该集合中按照如下贪婪规则选取一条边:

$$\min \frac{w'(e(v_i^x, v_j^y))}{\sum_{e(v_i^x, v_j^y) \in E_a} w'(e(v_i^x, v_j^y))} \quad (5.9),$$

选取这条边后, 将该边从 E_a 中删除, 并将与边 $e(v_i^x, v_j^y)$ 不兼容的边, 也即 $\bar{e}(v_i^x, v_j^y)$, 从 E_a 中删除; 与 $e(v_i^x, v_j^y)$ 不兼容的边是指, 比如: 当任务 v_i^x 选择在边缘执行, v_j^y 在云端执行, 那么与边 $e(v_i^E, v_j^C)$ 不兼容的边就是 $e(v_i^C, v_j^C)$, 该边代表任务 v_i^x, v_j^y 均选择在云端执行; $e(v_i^E, v_j^E)$, 该边代表任务 v_i^x, v_j^y 均选择在边缘执行; 以及 $e(v_i^C, v_j^E)$, 该边代表任务 v_i^x 选择在云端执行, v_j^y 在边缘执行。

接下来循环该步骤, 直至 E_a 为空。

以图5为例。第一步, 根据贪婪准则, 本算法选择了红色虚线边 (代表计算任务1在云端执行, 计算任务2卸载到边缘执行), 红色边代表与红色虚线边不兼容的边 (红色边分别代表了任务1在边缘端执行, 或者任务2在云端执行这几种情况, 与前面提到的计算任务1、2分别在云端和在边缘端执行相互矛盾)。将这些红色边以及红色虚线边同时从 E_a 中删除, 代表以后不会从这些边中选取下一条边。接下来如图6所示, 继续根据贪婪准则, 选取蓝色虚线边 (代表计算

任务2在边缘端执行，计算任务3在云端执行），与前面提到的相同，蓝色的边也代表了与蓝色虚线边不兼容的边。接下来同时将这些边从 E_a 中删除。由于 E_a 不为空，算法继续。接下来如图7所示，继续选取棕色虚线边，棕色虚线边代表了计算任务1在云端执行，计算任务4在边缘设备执行。然后将这些棕色虚线边以及实线边从 E_a 中删去。最后如图8所示，选取黄色的虚线边，代表计算任务1和计算任务3均在云端执行。将黄色边从 E_a 中删除后，由于 E_a 为空，算法停止。最后的卸载策略就是，计算任务1, 3在云端执行，计算任务2, 4在边缘设备上执行。

具体算法伪代码可以参见图1，算法中的符号及其含义对照表可以参见表2。

表1：算法伪代码

Algorithm 1: 边云计算新模型中的任务卸载算法	
1	输入: 计算任务图 $G(V, E)$: V 是点集合, 代表了计算任务
2	E 是边集合, 代表了任务之间的通信
3	输出: 计算任务卸载策略: V_C (在云端执行的任务)
4	V_E (在边缘端执行的任务)
5	C_{total} (计算任务执行总代价)
6	注: $x, y \in \{E(edge), C(cloud)\}$
7	for edge $e(v_i^x, v_j^y)$ do
8	$w'(e(v_i^x, v_j^y)) = w(e(v_i^x, v_j^y)) + \frac{w_i^x}{ N_i^x } + \frac{w_j^y}{ N_j^y };$
9	end
10	$E_a \leftarrow$ all edges in G
11	while $E_a \neq \emptyset$ do
12	for $e(v_i^x, v_j^y) \in E_a$ do
13	if $\frac{w'(e(v_i^x, v_j^y))}{\sum_{e(v_i^x, v_j^y) \in E_a} w'(e(v_i^x, v_j^y))}$ is minimal then
14	$V_x \leftarrow V_x \cup v_i$;
15	$V_y \leftarrow V_y \cup v_j$;
16	$E_a \leftarrow E_a \setminus e(v_i^x, v_j^y)$;
17	$E_a \leftarrow E_a \setminus \tilde{e}(v_i^x, v_j^y)$;
18	$C_{total} = C_{total} + w'(e(v_i^x, v_j^y))$;
19	end
20	end
21	end
22	return V_C, V_E, C_{total}

表2: 算法中的符号与含义对照表

符号	含义
V_C	在云端执行的任务
V_E	在边缘端执行的任务
E_a	计算任务图中的边集合
C_{total}	总执行代价 (计算代价 + 通信代价 + 任务传输代价)
$ N_i $	与任务 v_i 相互通信的任务的个数
x, y	x 和 y 分别代表云设备或者边缘设备
w_i^x	计算任务 v_i 在 x 端执行
$e(v_i^x, v_j^y)$	计算任务 v_i 在 x (边或者云) 端, v_j 在 y (边或者云) 端执行时, v_i 和 v_j 之间的代表通信的边.
$\bar{e}(v_i^x, v_j^y)$	与边 $e(v_i^x, v_j^y)$ 不兼容的边 (“不兼容” 的定义在权利点 5.3 中)
$w(e(v_i^x, v_j^y))$	计算任务 v_i 在 x (边或者云) 端, v_j 在 y (边或者云) 端执行时, v_i 和 v_j 之间的通信代价.

以上所述仅为本发明的实施例，并非以此限制本发明的保护范围，凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的系统领域，均同理包括在本发明的保护范围内。

1、一种基于代价优化的计算任务卸载算法，其特征在于，包括以下步骤：

1) 构建边云计算新模型；所述边云计算新模型包括计算三种重要的代价：

- 1.1) 计算任务的执行代价；
- 1.2) 同端计算任务之间的通信代价；
- 1.3) 跨端计算任务之间的非对称通信代价；
- 2) 对边云计算新模型进行扩展；
- 3) 合并计算代价；
- 4) 基于贪婪准则来求解优化卸载策略。

2、根据权利要求 1 所述的一种基于代价优化的计算任务卸载算法，其特征在于：上述步骤 1.1) 计算任务的执行代价，具体为：

给定任务图 G ，将任务的执行代价 $C_{comp}(G)$ 表达为如下形式：

$$C_{comp}(G) = \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \quad (5.1),$$

其中 w_i^E 和 w_i^C 分别表示任务 i 在边缘(E:edge)和在云端(C:cloud)的执行代价，布尔变量 $x_i = 0$ 的时候，代表任务 i 保留在云端执行， $x_i = 1$ 的时候，代表任务 i 卸载到边缘执行。

3、根据权利要求 2 所述的一种基于代价优化的计算任务卸载算法，其特征在于：上述步骤 1.2) 计算任务之间的通信代价，具体为：

用 $\langle w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4 \rangle$ 表示同端通信代价和跨端不对称的通信代价，也即任务 i 到任务 j 从边缘到边缘、边缘到云端、云端到边缘、边缘到边缘这四种通信代价；这四种通信代价统一用 $C_{comm}(G)$ 表示

为如下形式:

$$C_{comm}(G) = \sum_{i,j} (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) \quad (5.2)。$$

4、根据权利要求 3 所述的一种基于代价优化的计算任务卸载算法, 其特征在于: 上述步骤 1.3) 计算任务的卸载代价, 具体为:

任务 i 从云端卸载到边缘必须要考虑卸载代价, 任务卸载代价 $C_{tran}(G)$ 记为:

$$C_{tran}(G) = \sum_{i=1}^n m_i x_i \quad (5.3),$$

为了进一步简化, 将任务 i 的计算代价和卸载代价合并到一起:

$$\begin{aligned} C_{comp}(G) &= \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i + m_i x_i) \\ &= \sum_{i=1}^n ((w_i^E + m_i) x_i + w_i^C \bar{x}_i) \\ &= \sum_{i=1}^n ((w^E)_i x_i + w_i^C \bar{x}_i) \end{aligned} \quad (5.4),$$

因此, 原本总执行代价是计算代价、通信代价和卸载代价的总和, 因为合并了计算代价和卸载代价, 边缘计算中求解最优任务分配卸载可以表示为如下形式:

$$C_{total}(G) = \min_{x_i \in \{0,1\}} \{C_{comp}(G) + C_{comm}(G)\} \quad (5.5),$$

具体可表达为:

$$\begin{aligned} \min \{ & \sum_{i,j} (w_{ij}^1 x_i x_j + w_{ij}^2 x_i \bar{x}_j + w_{ij}^3 \bar{x}_i x_j + w_{ij}^4 \bar{x}_i \bar{x}_j) + \sum_{i=1}^n (w_i^E x_i + w_i^C \bar{x}_i) \} \\ \text{s.t. :} & \\ & x_i, x_j \in \{0,1\} \quad \forall i, j = 1, \dots, n. \end{aligned} \quad (5.6)。$$

5、根据权利要求 4 所述的一种基于代价优化的计算任务卸载算法, 其特征在于: 上述步骤 2) 中, 对边云计算新模型进行扩展具体

为;

边云计算新模型中, 每两个任务节点 v_i, v_j 之间的代表通信的边 $e(v_i, v_j)$ 上均有四种通信代价可能 $w_{ij}^1, w_{ij}^2, w_{ij}^3, w_{ij}^4$, 将四个权重的单条边扩展为四条边 $e(v_i^E, v_j^E)$ 、 $e(v_i^E, v_j^C)$ 、 $e(v_i^C, v_j^E)$ 和 $e(v_i^C, v_j^C)$, 每条边上一个确定的通信代价。

6、根据权利要求5所述的一种基于代价优化的计算任务卸载算法, 其特征在于: 上述步骤3) 中, 合并计算代价具体为:

将节点上的两个计算代价合并到通信代价中:

首先, 定义计算节点 v_i 的邻居节点 (与 v_i 存在通信的任务节点) 集合为 N_i :

$$N_i = \{v_j | e(v_i, v_j) = 1\} \quad (5.7),$$

对于每一条边 $e(v_i^x, v_j^y)$ ($x, y \in \{E(\text{Edge}), C(\text{Cloud})\}$), 将其边权 $w(e(v_i^x, v_j^y))$ 改为:

$$w'(e(v_i^x, v_j^y)) = w(e(v_i^x, v_j^y)) + \frac{w_i^x}{|N_i|} + \frac{w_j^y}{|N_j|} \quad (5.8)。$$

7、根据权利要求6所述的一种基于代价优化的计算任务卸载算法, 其特征在于: 上述步骤4) 中, 基于贪婪准则来求解优化卸载策略, 具体为:

定义 E_a 为所有边 $e(v_i^x, v_j^y)$ ($x, y \in \{E(\text{Edge}), C(\text{Cloud})\}$) 的集合, 每次从该集合中按照如下贪婪规则选取一条边:

$$\min \frac{w'(e(v_i^x, v_j^y))}{\sum_{e(v_i^x, v_j^y) \in E_a} w'(e(v_i^x, v_j^y))} \quad (5.9),$$

选取这条边后，将该边从 E_a 中删除，并将与边 $e(v_i^x, v_j^y)$ 不兼容的边，也即 $\bar{e}(v_i^x, v_j^y)$ ，从 E_a 中删除；

接下来循环该步骤，直至 E_a 为空。

8、根据权利要求7所述的一种基于代价优化的计算任务卸载算法，其特征在于：上述步骤4)中，

与 $e(v_i^x, v_j^y)$ 不兼容的边是指，比如：当任务 v_i^x 选择在边缘执行， v_j^y 在云端执行，那么与边 $e(v_i^E, v_j^C)$ 不兼容的边就是 $e(v_i^C, v_j^C)$ ，该边代表任务 v_i^x ， v_j^y 均选择在云端执行；

$e(v_i^E, v_j^E)$ ，该边代表任务 v_i^x ， v_j^y 均选择在边缘执行；以及 $e(v_i^C, v_j^E)$ ，该边代表任务 v_i^x 选择在云端执行， v_j^y 在边缘执行。

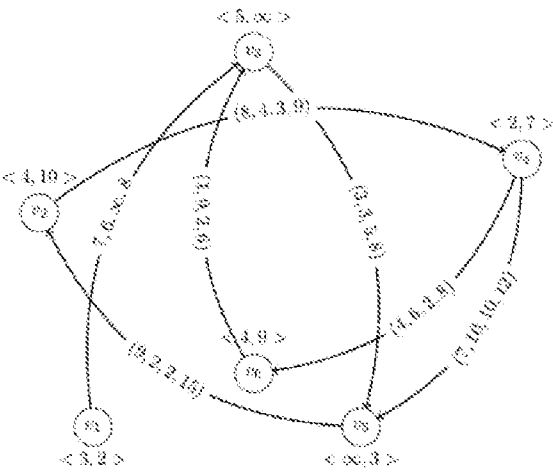


图 1

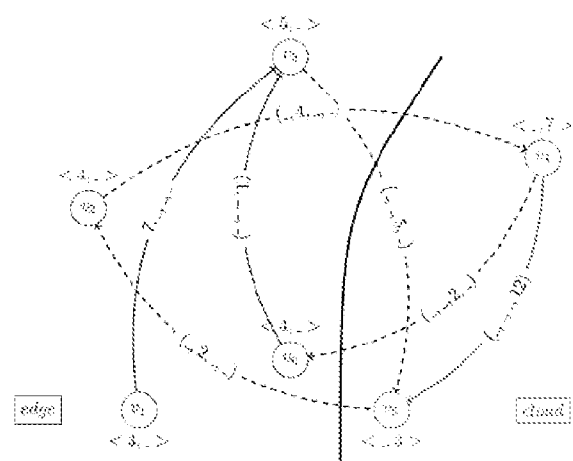


图 2

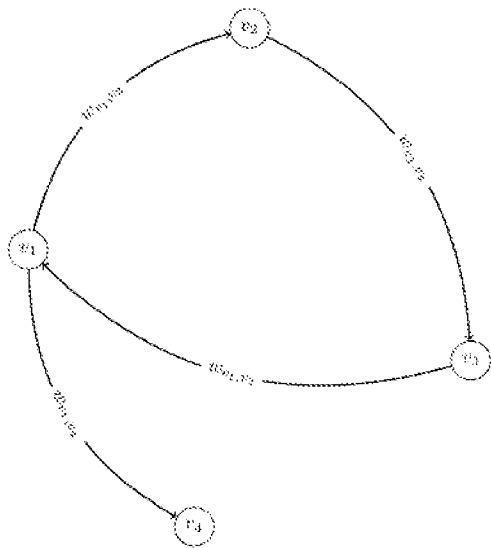


图 3

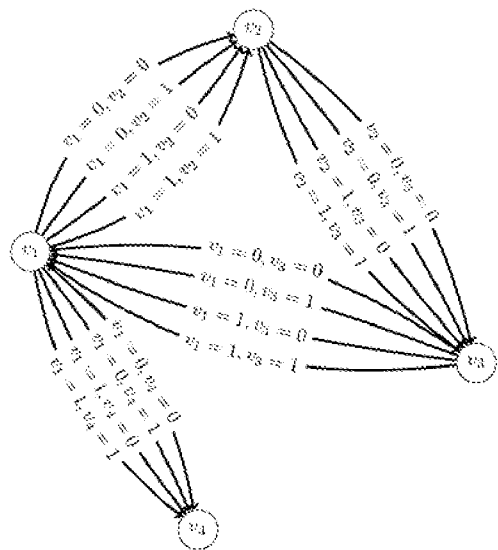


图 4

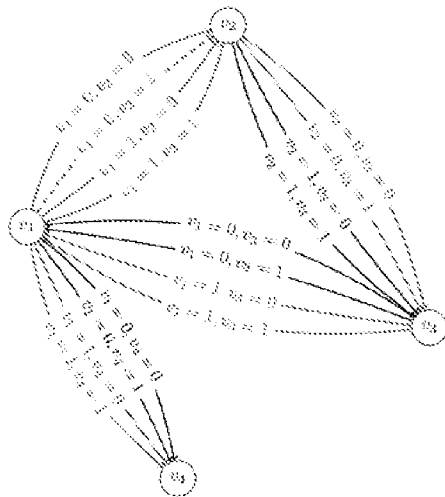


图 5

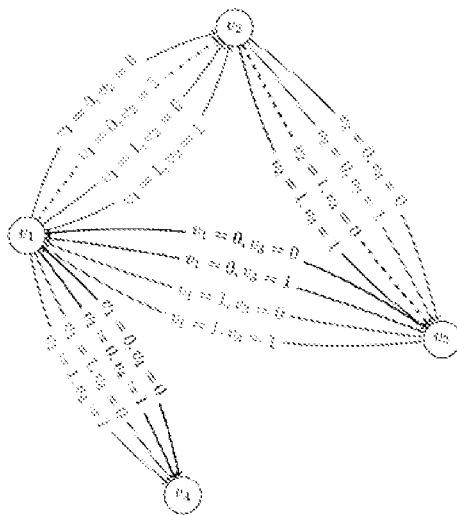


图 6

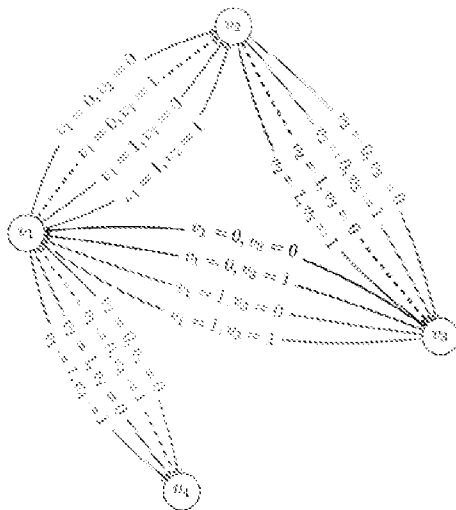


图 7

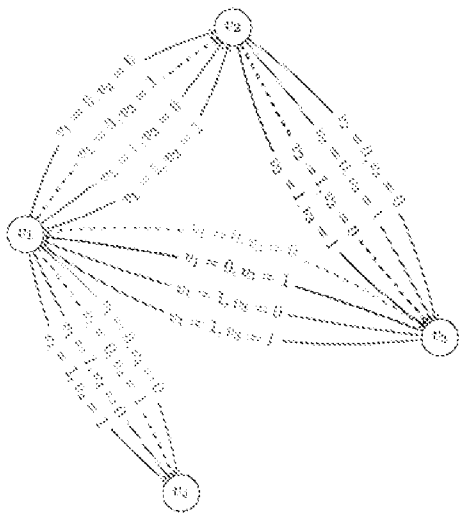


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/124084

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08(2006.01)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)

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)

CNPAT, CNKI, WPI, EPODOC: 云, 端, 计算, 卸载, 模型, 任务, 通信, cloud, edge, comput+al, unload, offload, uninstall, model, task, communicat+al, calculate.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109698861 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY) 30 April 2019 (2019-04-30) claims 1-8	1-8
A	CN 108600299 A (SHANGHAI INSTITUTE OF MICROSYSTEM AND INFORMATION TECHNOLOGY, CHINESE ACADEMY OF SCIENCES et al.) 28 September 2018 (2018-09-28) claims 1-8 and description, pp. 4-8	1-8
A	CN 104899103 A (NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY) 09 September 2015 (2015-09-09) entire document	1-8
A	CN 108920280 A (HARBIN INSTITUTE OF TECHNOLOGY) 30 November 2018 (2018-11-30) entire document	1-8
A	US 2017134186 A1 (MUKUNDAN, S. et al.) 11 May 2017 (2017-05-11) entire document	1-8



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

28 February 2020

Date of mailing of the international search report

06 March 2020

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
Information on patent family members

International application No.

PCT/CN2019/124084

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109698861	A	30 April 2019	None	
CN	108600299	A	28 September 2018	None	
CN	104899103	A	09 September 2015	None	
CN	108920280	A	30 November 2018	None	
US	2017134186	A1	11 May 2017	None	

国际检索报告

国际申请号

PCT/CN2019/124084

A. 主题的分类 H04L 29/08 (2006.01) 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) 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)) CNPAT, CNKI, WPI, EPDOC: 云, 端, 计算, 卸载, 模型, 任务, 通信, cloud, edge, comput+/al, unload, offload, uninstall, model, task, communicat+/al, calculate.		
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 109698861 A (深圳先进技术研究院) 2019年 4月 30日 (2019 - 04 - 30) 权利要求1-8	1-8
A	CN 108600299 A (中国科学院上海微系统与信息技术研究所等) 2018年 9月 28日 (2018 - 09 - 28) 权利要求1-8, 说明书第4-8页	1-8
A	CN 104899103 A (中国人民解放军国防科学技术大学) 2015年 9月 9日 (2015 - 09 - 09) 全文	1-8
A	CN 108920280 A (哈尔滨工业大学) 2018年 11月 30日 (2018 - 11 - 30) 全文	1-8
A	US 2017134186 A1 (MUKUNDAN SUNIL等) 2017年 5月 11日 (2017 - 05 - 11) 全文	1-8
☐ 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 2020年 2月 28日		国际检索报告邮寄日期 2020年 3月 6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 李彬 电话号码 53961633

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/124084

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	109698861	A	2019年 4月 30日	无	
CN	108600299	A	2018年 9月 28日	无	
CN	104899103	A	2015年 9月 9日	无	
CN	108920280	A	2018年 11月 30日	无	
US	2017134186	A1	2017年 5月 11日	无	