高空平台站（HAPS）辅助车联网中的切换感知型分布式计算

Qiqi Ren、Omid Abbasi（IEEE 高级会员）、Gunes Karabulut Kurt（IEEE 高级会员）、Halim Yanikomeroglu（IEEE 会士）和 Jian Chen（IEEE 会员）

arXiv:2305.04196v1 [cs.DC] 7 May 2023

摘要—分布式计算通过协同利用网络边缘和车辆的计算资源，为车联网（IoV）服务提供支持。然而，频繁的边缘网络切换会导致计算中断，计算资源严重短缺也是提供车联网服务时面临的两个问题。高空平台站计算凭借其广泛的覆盖范围和强大的计算能力，可作为现有分布式计算框架的有力补充。为此，本文在一种包含高空平台站计算的新型分布式计算框架中提出了一种自适应方案，以解决车联网面临的这两个问题。根据车辆的不同需求、网络动态性和切换时间敏感性，所提方案灵活地将每个任务划分为三部分，并将其分配给车辆、路侧单元（RSU）和高空平台站执行同步计算。该方案还限制路侧单元上的任务计算，使其在切换前完成，以避免计算中断的风险。在此基础上，我们构建了一个考虑任务分割比例、发射功率、带宽分配和计算资源分配的时延最小化问题。为解决该问题，我们提出了一种基于变量替换和逐次凸近似的方法。仿真结果表明，该方案不仅能够灵活避免切换带来的负面影响，还兼顾了时延性能并保持了时延稳定性。

索引词——车联网（IoV）、高空平台站（HAPS）、分布式计算、边缘网络切换。

I. 引言

A. 背景

随着无线通信、物联网和人工智能（AI）等技术的发展，车联网（IoV）这一概念也逐渐受到关注发展势头强劲，自动驾驶被视为最终目标 [1], [2]。作为车联网的主要组成部分，智能网联汽车（ICV）将整合感知、决策和控制功能。未来，智能网联车辆（ICVs）将实现从物理空间到信息空间的大量信息转换，而这一转换将依赖细粒度数据计算，包括各种基于机器学习的训练和推理模型 [3], [4]。过去几年中，机器学习领域已从集中式云处理的大数据范式，转向由网络边缘设备进行分布式处理的小数据范式 [5], [6]。这一转变的推动力在于同步利用本地可用资源和边缘节点的邻近资源，以获得基于人工智能的任务计算的实时响应，从而促进分布式计算的发展 [7], [8]。鉴于分布式计算具有邻近性、效率、可扩展性和易于协作等优势，预计其将在实现自动驾驶方面发挥重要作用。

自动驾驶预计每天会产生太字节级数据 [9]。在此背景下，未来的车联网可能持续面临计算资源短缺的问题，而且这一短缺情况会比普通用户网络更加严重。这一预测有两个原因：首先，与仅在使用时才占用资源的普通用户不同，智能网联车辆通常长时间在线，因此会持续占用计算资源。其次，车联网服务对安全性的要求高于普通用户服务。提供优质的车联网服务直接关系到道路安全和人身安全，这意味着计算资源短缺将带来严重的安全问题 [10]。为解决这一问题，人们提出了多种方案，包括车辆协同计算、无人机计算和卫星计算，以补充边缘计算并构建新的分布式计算框架 [11]–[14]。然而，车车连接中断、无人机驻留时间短且不可靠、卫星传输时延过长，以及复杂的移动性管理，使得这些方案难以满足车联网在安全性、稳定性和可靠性方面的需求。因此，我们开始探索其他更可行的解决方案。近年来，高空平台站（HAPS）系统被提出作为6G网络的候选方案 [15], [16]。高空平台站是一种空中平台，例如

本研究部分得到国家自然科学基金（项目编号：61771366）的资助，部分得到华为加拿大的资助（通讯作者：Jian Chen）。

Q. Ren 任职于中国陕西省西安市 710071 的西安电子科技大学综合业务网络国家重点实验室，以及加拿大安大略省渥太华市卡尔顿大学系统与计算机工程系非地面网络（NTN）实验室（电子邮箱：renqiqi5277@gmail.com）。

J. Chen 任职于中国陕西省西安市 710071 的西安电子科技大学综合业务网国家重点实验室（电子邮箱：jianchen@mail.xidian.edu.cn）。

O. Abbasi 和 H. Yanikomeroglu 任职于加拿大安大略省渥太华市卡尔顿大学系统与计算机工程系非地面网络（NTN）实验室（电子邮箱：{omidabbasi, halim}@sce.carleton.ca）。

G. Karabulut Kurt 任职于加拿大蒙特利尔的蒙特利尔综合理工学院电气工程系（电子邮箱：gunes.kurt@polymtl.ca）

飞艇或气球，可作为无线通信站长期部署于平流层，从高空俯瞰大范围地面区域（半径为50-500 km），并提供视距（LoS）覆盖 [17]。由于高空平台站具有较大的有效载荷（通常为 $\geq$ 100 kg），因此能够搭载多种资源（天线、容量、计算资源、存储等），为地面计算系统提供强有力的补充 [18], [19]。此外，当前和未来的太阳能、风能转换技术以及电池技术，可为高空平台站提供强大的能源供应潜力 [20]–[22]。大量研究已证明，将高空平台站计算引入物联网网络可在能源和时延性能方面带来收益 [23], [24]。受此启发，我们在近期的研究中将高空平台站计算引入车联网 [25] ，并提出了一种计算卸载方案，以加速车联网服务。

除了计算资源严重短缺的前景外，未来的车联网还面临网络切换问题。在实际应用中，边缘节点的通信覆盖范围通常较小（不到 300 m [26]）。为确保网络连接和服务的可持续性，智能网联车辆在跨越不同边缘节点的覆盖范围时，必然会触发网络连接切换。然而，发生网络切换时，智能网联汽车可能会经历卸载过程的暂时中断，或者由于超出计算目标节点的覆盖范围而无法直接接收计算结果。为解决这一问题，研究人员提出了一系列计算卸载方案，以应对切换问题 [27]–[32]。 [27], [28] 中的作者考虑了计算性能指标，以设计切换方案。 [27], 研究了如何在最大限度延长车辆与边缘服务器连接时间的同时，尽可能减少切换次数。 [28],在考虑边缘服务器资源短缺的同时，基于联合优化网络切换和任务迁移决策，利用远程云计算资源协助处理连接的边缘服务器无法及时处理的任务。然而，车联网切换问题会受到多种因素影响，包括接收信号强度、车辆速度、可靠性和效率。此外，在制定切换决策时考虑多维因素，会使问题求解过程更加复杂 [33]–[35]。对此， [29]–[32] 设计了计算迁移方案，以适应网络切换并消除其对计算性能的负面影响。因此， [29] 提出了一种在发生切换时于多台服务器之间迁移任务的方案。 [30] 为任务设置了“挂起”模式，当任务无法在连接到下一台可用服务器之前处理完毕时，车辆将暂停任务调度。鉴于切换可能导致任务结果无法送达车辆，张H等人提出在多台服务器之间迁移结果 [31]。基于路径规划可预测这一假设， [32] 提出对计算任务和结果进行预测性上传和下载，以提高结果获取效率。

B. 动机与贡献

从现有方案来看，分布式计算中适应边缘网络切换的主流方法是在边缘节点之间进行计算迁移。然而，这些方法缺乏灵活性，且其性能无法得到保证。具体而言，大多数方法都提出对受切换影响的计算任务采取补偿措施，例如任务和结果迁移、任务重新卸载以及任务延后。虽然这些设计旨在确保任务得以执行或结果成功送达，但此类方法具有被动性且缺乏灵活性。此外，由于只考虑边缘节点资源，任务或其结果通常需要在多个边缘节点之间迁移，这会导致响应时间过长，甚至无法响应，从而造成性能大幅波动，甚至计算失败。因此，为应对切换，计算卸载方案需要避免切换对计算造成的负面影响，同时确保计算任务在切换和未发生切换的情况下都具有低延迟性能，从而为智能网联汽车提供流畅、稳定的服务体验。

幸运的是，高空平台站（HAPS）能够帮助解决上述问题。凭借其对地面的俯瞰视角和强大的计算资源，高空平台站能够应对非典型的需求激增，并避免边缘网络切换引起的计算中断。因此，将高空平台站计算引入车联网系统，既能缓解计算资源短缺问题，又能有效应对切换问题，这对于智能网联汽车而言具有重要意义。基于这一关键观察，我们遵循近期工作 [25] 中提出的分布式计算框架。在此基础上，我们提出了一种切换感知计算卸载方案，以解决计算中断问题。在该方案中，任务将被动态拆分，并分布式分配给智能网联汽车、路侧单元和高空平台系统节点进行同步处理。这意味着，在这三类节点的协同作用下，智能网联汽车的任务能够通过并行方式得到处理。因此，当即将发生网络切换时，路侧单元仅负责处理极少量的数据，以确保分配给该节点的任务部分能够在切换发生前完成计算，而高空平台站则承担大部分计算负担，从而顺利完成任务计算。本研究的贡献如下：

1) 我们利用一种整合智能网联车辆、路侧单元和高空平台节点计算资源的分布式计算框架，提出了一种切换感知计算方案。与现有边缘网络解决方案不同，该方案利用高空平台站辅助分布式框架，灵活应对边缘网络切换导致的计算中断问题。据我们所知，这是首项研究高空平台站计算在这一主题上的优势的工作。[1]

[1]在本研究中，我们沿用前期工作中的集成三层计算框架 [25]，，并提出一种新的计算卸载方案，以应对边缘网络切换造成的计算中断问题。

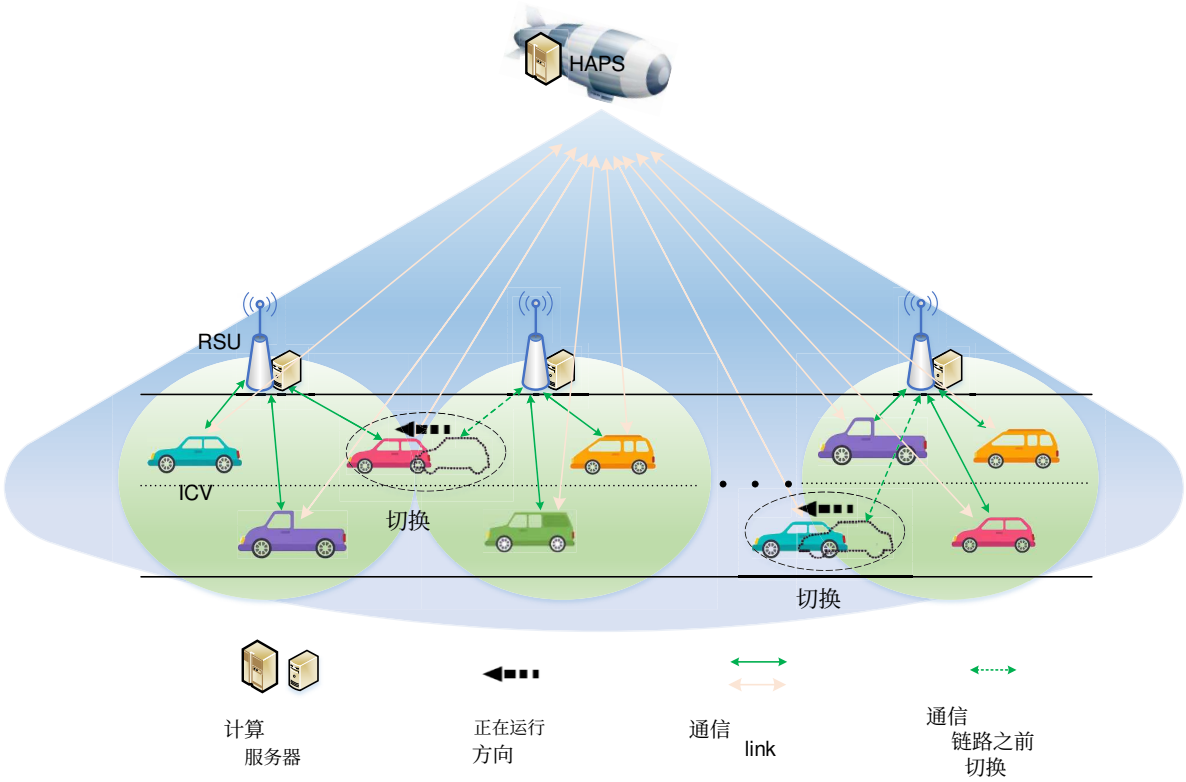


图 1：高空平台辅助车联网中的切换感知分布式计算。

2) 鉴于智能网联车辆需求多样、网络动态性以及切换时间敏感性，该方案灵活地将每个用户的任务划分为三部分，并分配给智能网联车辆、路侧单元和高空平台节点进行同步计算。此外，该方案对路侧单元上的任务计算加以限制，确保其在即将发生切换之前完成，从而避免计算中断的风险。

3）为了最小化智能网联车辆任务的执行时延，我们通过求解任务拆分比例、智能网联汽车发射功率、带宽分配和计算资源分配，构建了一个优化问题。随后，我们使用变量替换和逐次凸近似（SCA）方法求解该优化问题。

II. 切换感知分布式计算模型

A. 系统模型

本文所考虑的高空平台辅助车联网切换感知分布式计算模型如图1所示。单个高空平台站位于20 km高度，为一条单向道路 [36]提供覆盖。该道路被划分为 M 个互不重叠的路段，每个路段均由一个边缘网络接入点（即路侧单元）覆盖。每个路侧单元都配备一台计算服务器，用于处理车辆计算任务；路侧单元及其配备的计算资源均可视为系统中的边缘网络资源。这些路侧单元依次标记为 $\mathcal{M}=\{1,2,\ldots,M\}$。我们假设其覆盖范围为 D (m)。有 N 辆智能网联汽车，依次标记为 $\mathcal{N}=\{1,2,\ldots,N\}$，以 v(m/s)的速度沿道路行驶，且每辆智能网联汽车都会依次经过各路侧单元的覆盖范围。这意味着，智能网联汽车从一个路侧单元的覆盖区域驶入另一个路侧单元的覆盖区域时，就会发生切换。在每个决策时隙开始时，智能网联汽车 n 从当前位置 l_n (m)出发，并生成一个待计算任务，其输入数据量为 ε_n (比特)，计算密度为λ_n (CPU周期/比特)。假设任务可拆分，我们考虑图2所示的分布式并行计算模型。在该模型中，任务可拆分为三个部分，并在三个节点上同时计算（即在智能网联汽车的车载设备、关联路侧单元和高空平台站上计算），对应三种计算方式：本地计算、路侧单元计算和高空平台站计算。本地计算可直接处理分配给它的任务部分，而分配给路侧单元和高空平台站的任务部分，则需分别卸载至相应节点，之后由其服务器执行计算。本文不讨论结果下载，因为与输入数据相比，这类数据量较小。鉴于时延性能对车联网服务至关重要，且会影响车辆与道路安全以及驾驶员体验，本文致力于提升智能网联汽车执行计算任务时的时延性能。

根据图2所示模型，数据划分模型描述如下：对于智能网联汽车 n，其关联路侧单元节点处的数据划分比例表示为 x_n^R，在

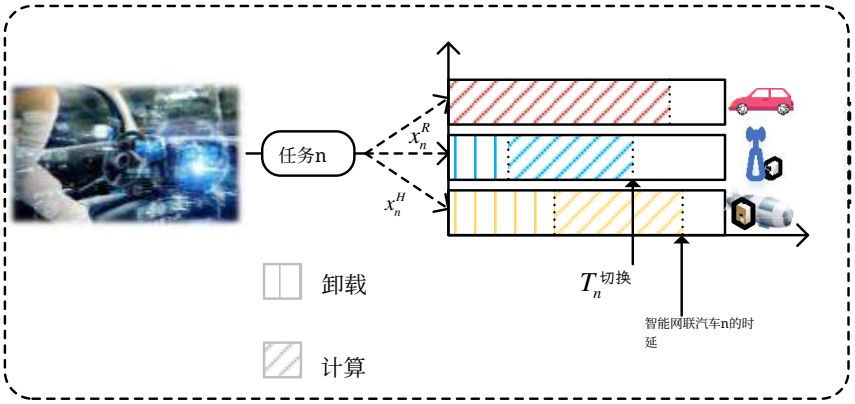


图2：分布式并行计算模型。

高空平台系统节点处表示为 x_n^H，在其车载设备处表示为 $1 - x_n^R - x_n^H$。任务划分策略是动态确定的，而非采用固定分配方式。该策略取决于实时信道条件、用户与路侧单元之间的关联关系、变化的任务需求（数据量和计算密度），以及可用带宽和计算资源容量 [37]。除了这些因素，还通过灵活调整数据划分比例来考虑边缘网络的切换问题，以确保路侧单元上计算的任务部分不会因切换而中断。图3展示了切换过程。如图所示，在一个决策时隙开始时，智能网联汽车 n 从位置 A 移动至网络边界 B，即将进入下一个路侧单元的覆盖范围，从而触发网络切换。根据智能网联汽车的初始速度 v（米/秒）及其位置 l_n，切换时间戳可表示为 $T_n^{handoff} = {}^{D-|l_n}{}_{取模\ v}{}^{D|}$，其中 $|l_n$ 取模 $D|$表示 l_n 除以 D所得的余数，该值表示智能网联汽车 n相对于其关联路侧~~单元的位置~~。[2] $T_n^{handoff}$描述智能网联汽车 n 驶出路侧单元覆盖边界所需的时间。在此过程中，智能网联汽车 n 以划分比例 x_n^R 将部分任务卸载至路侧单元，随后由路侧单元服务器计算分配的任务部分。为确保数据处理不因网络切换而中断，所经历的时延必须保证不超过切换发生的时间戳，即 $T_n^{handoff}$，这可视为硬时间限制。如果路侧单元上该任务部分的处理时延大于 $T_n^{handoff}$，则相应部分将被丢弃。

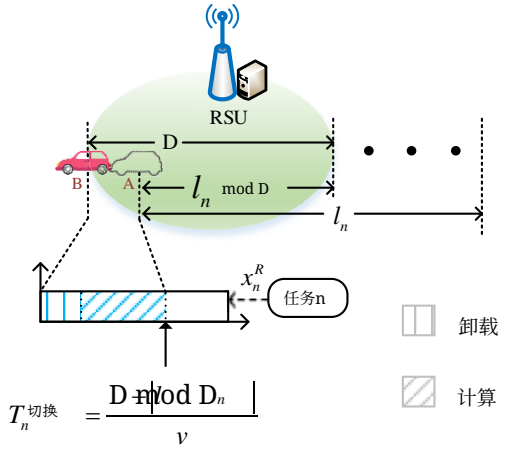


图3：切换示意图。

B. 通信模型

在该系统中，我们假设路侧单元和智能网联车辆均配备单天线，而高空平台站配备多天线。基于这一假设，高空平台站网络中有大量点波束，并且目标区域位于一个点波束 [25]内。在每个时间槽，智能网联汽车 n 需要将大小为 $x_n^R \varepsilon_n$ 的任务部分的输入数据卸载到关联的路侧单元，同时将大小为 $x_n^H \varepsilon_n$ 的任务部分的输入数据卸载到高空平台站。因此，通信链路分为两种：智能网联车辆与路侧单元之间的链路，以及智能网联车辆与高空平台站之间的链路。我们为所有链路考虑 2 GHz 的载波频率，且这些链路使用正交信道。给定可用带宽 $B_{\max}$，将对这些正交链路的带宽进行最优分配。对于智能网联汽车 n，分配给该车至路侧单元链路和至高空平台站链路的带宽比分别表示为 b_n^R和 b_n^H。带宽分配应满足紧致性约束 $\sum b_n^R + b_n^H \leq 1$。此外，考虑到智能网联车辆发射功率的限制，我们还将研究给定 $P_{\max}$时的功率分配。对于智能网联汽车n，该车 n 至路侧单元链路和该车 n 至高空平台站链路的功率分配比分别表示为 p_n^R和 p_n^H。此处，我们使用 $p_n^R + p_n^H \leq 1$、$\forall n \in \mathcal{N}$来

[2]该模型根据基于距离的切换协议规则进行构建。

表 I：符号定义

符号说明	定义
$\mathcal{M}, M$	路侧单元集合及路侧单元数量
$\mathcal{N}, N$	智能网联汽车集合及智能网联车辆数量
ε	输入数据量（Kbits）
λ	计算密度（CPU周期/比特）
x_n^R, x_n^H	路侧单元节点和高空平台系统节点处的数据分割比例
D	路侧单元的覆盖范围（米）
F^L, F^R, F^H	智能网联汽车、路侧单元和高空平台站的计算能力（CPU周期/秒）
G^R, G^H, G	RSU 链路信道增益、HAPS 链路信道增益和定向天线增益
d^R, d^H	RSU 和 HAPS 的链路距离（m）
c, f_c, α, β_0	光速（m/s）、载波频率（Hz）、非视距链路路径损耗因子和非视距链路参考路径损耗
h^R, h^H	视距链路小尺度衰落系数和非视距链路小尺度衰落系数
N_0	高斯噪声功率谱密度（dBm/Hz）
$B_{\max}, R^R, R^H$	带宽限制（兆赫）、智能网联汽车到路侧单元的传输速率，以及智能网联汽车到高空平台站的传输速率（比特/秒）
f, b	计算资源分配比例和带宽分配比例
$P_{\max}$	智能网联汽车的发射机功率限制（分贝毫瓦）
T^L, T^R, T^H	本地、路侧单元和高空平台站计算下的时延（秒）
$T_n^{handoff}$	触发切换的时间（秒）

约束智能网联车辆的功率分配。我们考虑智能网联汽车 n至路侧单元链路上的非视距（NLoS）通信，并将信道增益建模为

$$G_n^R = \frac{\beta_0(f_c)\left|h_n^R\right|^2}{(d_n^R)^\alpha}, \tag{1}$$

其中， β_0（f_c）表示参考距离 1 m 处的路径损耗， h_n^R 表示服从瑞利分布的非视距链路小尺度衰落系数， d_n^R 表示智能网联汽车 n 与路侧单元之间的距离， α 表示路径损耗指数。相应的传输速率为 $R_n^R = b_n^R B_{\max}\log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right)$，其中 N_0 表示高斯噪声功率谱密度。

由于高空平台站（HAPS）悬停在高空，且高速公路通常位于偏远地区，因此我们认为智能网联车辆（ICVs）与高空平台站（HAPS）之间的链路均为视距（LoS）链路，其大尺度衰落路径损耗遵循自由空间路径损耗模型。信道增益可建模为 [38]:

$$G_n^H = G\left(\frac{c}{4\pi d_n^H f_c}\right)^2\left|h_n^H\right|^2, \tag{2}$$

其中， c 为光速， d_n^H 为智能网联汽车 n与高空平台站之间的距离。 f_c 为载波频率，在本文中取值为 2 GHz。由于 10 GHz 以下频率的环境影响可以忽略，因此本文不考虑环境衰减 [39]。 G 为定向天线增益， h_n^H 为考虑视距分量的莱斯衰落小尺度衰落系数。相应的传输速率为 $R_n^H = b_n^H B_{\max}\log_2\left(1+\frac{p_n^H P_{\max} G_n^H}{b_n^H B_{\max} N_0}\right)$。根据该传输速率，将智能网联汽车 n 大小为 $x_n^R\varepsilon_n$ 的任务部分卸载至路侧单元的时延可建模为 $\frac{x_n^R\varepsilon_n}{R_n^R}$，将大小为 $x_n^H\varepsilon_n$ 的智能网联汽车 n 任务部分卸载至高空平台站的时延可建模为 $\frac{x_n^H\varepsilon_n}{R_n^H}$。此外，智能网联汽车信号到达路侧单元和高空平台站所需的传播时延分别为 $\frac{d_n^R}{c}$和 $\frac{d_n^H}{c}$。

C. 计算模型

假设在每个决策时隙中，每辆智能网联汽车都会生成一个待计算任务，该任务可拆分为三个部分，并由三个节点并行计算。智能网联汽车进行本地计算、路侧单元计算和高空平台站计算时可用的计算资源分别为 F^L、 F^R 和 F^H（单位：CPU周期/秒）。拆分后，分配给路侧单元、工作负载为 $x_n^R\varepsilon_n\lambda_n$ 的任务部分将在路侧单元服务器上计算；分配给高空平台站、工作负载为$x_n^H\varepsilon_n\lambda_n$ 的任务部分将在高空平台站上计算；分配给智能网联汽车、工作负载为 $(1-x_n^R-x_n^H)\varepsilon_n\lambda_n$ 的任务部分将在本地计算。路侧单元服务器和高空平台站服务器的计算资源将进行优化。对于路侧单元计算，计算资源将在路侧单元覆盖范围内的智能网联汽车之间分配。令 Φ_m 表示路侧单元 m 的智能网联汽车索引集合， $m\in\mathcal{M}$。令 f_n^R 表示其关联路侧单元服务器分配给智能网联汽车 n 的计算资源比例，则路侧单元 m 的计算资源分配限制为 $\sum_{n\in\Phi_m} f_n^R \le 1$、$\forall m\in\mathcal{M}$。令 f_n^H 表示高空平台站服务器分配给智能网联汽车 n 的计算资源比例，高空平台站的计算资源分配限制为 $\sum_{n\in\mathcal{N}} f_n^H \le 1$。此外，本地计算、路侧单元计算和高空平台站计算的计算时延分别可表示为 $\frac{(1-x_n^R-x_n^H)\varepsilon_n\lambda_n}{F^L}$、 $\frac{x_n^R\varepsilon_n\lambda_n}{f_n^R F^R}$ 和 $\frac{x_n^H\varepsilon_n\lambda_n}{f_n^H F^H}$。

D. 时延模型

基于上述通信和计算模型，智能网联汽车 n 采用本地、路侧单元和高空平台站计算时的时延如下:

1）本地计算:

$$T_n^L = \frac{(1-x_n^R-x_n^H)\varepsilon_n\lambda_n}{F^L}. \tag{3}$$

2）路侧单元计算:

$$T_n^R = \frac{d_n^R}{c} + \frac{x_n^R\varepsilon_n}{b_n^R B_{\max}\log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right)} + \frac{x_n^R\varepsilon_n\lambda_n}{f_n^R F^R}. \tag{4}$$

3）高空平台站计算：

$$T_n^H = \frac{d_n^H}{c} + \frac{x_n^H \varepsilon_n}{b_n^H B_{\max} \log_2\left(1+\frac{p_n^H P_{\max} G_n^H}{b_n^H B_{\max} N_0}\right)} + \frac{x_n^H \varepsilon_n \lambda_n}{f_n^H F^H}. \quad (5)$$

智能网联汽车 n 的时延最终由 T_n^L、T_n^R 和 T_n^H [3] 中的最大值决定。

III. 总时延最小化

本节首先构建优化问题，以最小化所有智能网联车辆的总时延，从而提升系统的平均时延性能。随后，我们针对所构建的问题给出问题转换与求解方案。

A. 问题建模

我们的目标是通过求得 RSU 的任务拆分比例 $\mathbf{X}^{\boldsymbol{R}} = [x_1^R, ..., x_N^R]$、HAPS 的任务拆分比例 $\mathbf{X}^{\boldsymbol{H}} = [x_1^H, ..., x_N^H]$、ICV 到 RSU 链路的带宽分配 $\mathbf{B}^{\boldsymbol{R}} = [b_1^R, ..., b_N^R]$、ICV 到 HAPS 链路的带宽分配 $\mathbf{B}^{\boldsymbol{H}} = [b_1^H, ..., b_N^H]$、ICV 到 RSU 链路的功率分配 $\mathbf{P}^{\boldsymbol{R}} = [p_1^R, ..., p_N^R]$、ICV 到 HAPS 链路的功率分配 $\mathbf{P}^{\boldsymbol{H}} = [p_1^H, ..., p_N^H]$、RSU 的计算资源分配 $\mathbf{F}^{\boldsymbol{R}} = [f_1^R, ..., f_N^R]$ 以及 HAPS 的计算资源分配 $\mathbf{F}^{\boldsymbol{H}} = [f_1^H, ..., f_N^H]$ 的最优值，来优化 ICV 执行任务所经历的总时延。该优化问题可表述如下：

$$\min_{\boldsymbol{x}^R, \boldsymbol{x}^H, \boldsymbol{b}^R, \boldsymbol{b}^H, \boldsymbol{p}^R, \boldsymbol{p}^H, \boldsymbol{f}^R, \boldsymbol{f}^H} \quad \sum_n^N \max \quad \{T_n^L, T_n^R, T_n^H\} \quad (6a)$$

$$\text{s.t.} \quad T_n^R \le T_n^{handoff}, \quad \forall n \in \mathcal{N}, \quad (6b)$$

$$\sum_{n\in\mathcal{N}} b_n^R + b_n^H \le 1, \quad (6c)$$

$$p_n^R + p_n^H \le 1, \ \forall n \in \mathcal{N}, \quad (6d)$$

$$\sum_{n\in\Phi_m} f_n^R \le 1, \ \forall m \in \mathcal{M}, \quad (6e)$$

$$\sum_{n\in\mathcal{N}} f_n^H \le 1, \quad (6f)$$

$$x_n^R, x_n^H, b_n^R, b_n^H, p_n^R, p_n^H, f_n^R, f_n^H > 0. \quad (6g)$$

由于我们考虑的是并行任务计算，智能网联汽车（ICV）的时延是指在本地、路侧单元（RSU）和高空平台站（HAPS）上执行任务的时延中的最大值，即 $\max \quad \{T_n^L, T_n^R, T_n^H\}$。在问题 (6) 中，目标函数 (6a) 表示智能网联车辆的总时延。约束 (6b) 表示分配给路侧单元（RSU）的任务部分应在切换发生之前完成。(6c) 表示所有链路的带宽约束。(6d) 表示每个对象的发射功率约束。

[3]根据 $T_c = \lambda/v$，其中 λ 为波长，v 为车辆速度，我们场景中的相干时间为 5 毫秒。在相干时间内，仅分配 66.6 微秒用于信道估计（假设子载波为 15 KHz，且每个导频信号分配一个符号），剩余时间可用于卸载。这意味着，与卸载时间相比，信道估计时间可以忽略不计。

智能网联汽车。(6e) 表示每个路侧单元的计算资源约束，(6f) 表示高空平台站的计算资源约束。最后，(6g) 表示所有变量的非负要求。显然，由于目标函数非光滑、变量之间存在耦合关系，且数学表达式复杂，问题 (6) 是一个非凸问题。该问题难以直接求解，目前也没有标准方法能够解决这一问题。接下来，我们将该问题转化为一个易于求解的问题。

B. 问题转换与求解

首先，我们引入辅助变量 $\mathbf{T} = [T_1, ..., T_N]$ 来表示单个时延，从而将原始的非光滑问题转化为光滑问题。此外，通过展开 T_n^L、T_n^R 和 T_n^H 的表达式，问题 (6) 可以等价地重写为

$$\min_{\mathbf{T}, \boldsymbol{x}^R, \boldsymbol{x}^H, \boldsymbol{b}^R, \boldsymbol{b}^H, \boldsymbol{p}^R, \boldsymbol{p}^H, \boldsymbol{f}^R, \boldsymbol{f}^H} \quad \sum_n^N \quad T_n \quad (7a)$$

$$\text{s.t.} \quad \frac{(1-x_n^R-x_n^H)\varepsilon_n\lambda_n}{F^L} \le T_n, \ \forall n \in \mathcal{N}, \quad (7b)$$

$$\frac{d_n^R}{c} + \frac{x_n^R \varepsilon_n}{b_n^R B_{\max} \log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right)} + \frac{x_n^R \varepsilon_n \lambda_n}{f_n^R F^R} \le T_n, \quad \forall n \in \mathcal{N}, \quad (7c)$$

$$\frac{d_n^H}{c} + \frac{x_n^H \varepsilon_n}{b_n^H B_{\max} \log_2\left(1+\frac{p_n^H P_{\max} G_n^H}{b_n^H B_{\max} N_0}\right)} + \frac{x_n^H \varepsilon_n \lambda_n}{f_n^H F^H} \le T_n, \quad \forall n \in \mathcal{N}, \quad (7d)$$

$$\frac{d_n^R}{c} + \frac{x_n^R \varepsilon_n}{b_n^R B_{\max} \log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right)} + \frac{x_n^R \varepsilon_n \lambda_n}{f_n^R F^R} \le T_n^{handoff}, \quad \forall n \in \mathcal{N}, \quad (7e)$$

$$(6c), (6d), (6e), (6f), (6g). \quad (7f)$$

接下来，为了处理约束 (7c)、(7d) 和 (7e) 中优化变量之间的耦合关系，我们引入辅助变量 $\boldsymbol{\tau}^{\boldsymbol{R}} = [\tau_1^R, ..., \tau_N^R]$ 和 $\boldsymbol{\tau}^{\boldsymbol{H}} = [\tau_1^H, ..., \tau_N^H]$，分别表示智能网联车辆将数据卸载到路侧单元和高空平台站所产生的时延，并放宽相应的约束。于是，问题 (7) 可以等价地改写为

$$\min_{\mathbf{T}, \boldsymbol{x}^R, \boldsymbol{x}^H, \boldsymbol{b}^R, \boldsymbol{b}^H, \boldsymbol{p}^R, \boldsymbol{p}^H, \boldsymbol{\tau}^R, \boldsymbol{\tau}^H, \boldsymbol{f}^R, \boldsymbol{f}^H} \quad \sum_{\mathrm{n}}^N T_n \quad (8a)$$

$$\text{s.t.} \quad \frac{d_n^R}{c} + \tau_n^R + \frac{x_n^R \varepsilon_n \lambda_n}{f_n^R F^R} \le T_n, \ \forall n \in \mathcal{N}, \quad (8b)$$

$$\frac{d_n^H}{c} + \tau_n^H + \frac{x_n^H \varepsilon_n \lambda_n}{f_n^H F^H} \le T_n, \ \forall n \in \mathcal{N}, \quad (8c)$$

$$\frac{x_n^R \varepsilon_n}{b_n^R B_{\max} \log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right)} \le \tau_n^R, \ \forall n \in \mathcal{N}, \quad (8d)$$

$$\frac{x_n^H \varepsilon_n}{b_n^H B_{\max} \log_2\left(1+\frac{p_n^H P_{\max} G_n^H}{b_n^H B_{\max} N_0}\right)} \le \tau_n^H,\ \forall n \in \mathcal{N}, \tag{8e}$$

$$\frac{d_n^R}{c} + \tau_n^R + \frac{x_n^R \varepsilon_n \lambda_n}{f_n^R F^R} \le T_n^{handoff},\ \forall n \in \mathcal{N}, \tag{8f}$$

$$(6c)-(6f), (7b),$$

$$x_n^R, x_n^H, b_n^R, b_n^H, p_n^R, p_n^H, \tau_n^R, \tau_n^H, f_n^R, f_n^H > 0, \tag{8g}$$

其中，(8d) 确保智能网联车辆将数据卸载到路侧单元所经历的通信时延不超过 $\boldsymbol{\tau^R}$，而 (8e) 确保智能网联车辆将数据卸载到高空平台站所经历的通信时延不超过 $\boldsymbol{\tau^H}$。然后，为了进一步解决变量耦合问题，我们对变量 $\boldsymbol{x^R}$、$\boldsymbol{x^H}$、$\boldsymbol{\tau^R}$、$\boldsymbol{\tau^H}$、$\boldsymbol{f^R}$和 $\boldsymbol{f^H}$进行指数变换。具体而言，智能网联车辆 n 对应的上述变量可转换如下：$x_n^R = \exp(\overline{x_n^R})$ $x_n^H = \exp(\overline{x_n^H})$ $\tau_n^R = \exp(\overline{\tau_n^R})$ $\tau_n^H = \exp(\overline{\tau_n^H})$ $f_n^R = \exp(\overline{f_n^R})$ 和 $f_n^H = \exp(\overline{f_n^H})$。由此可得到以下优化问题：

$$\min_{\mathrm{T}, \overline{\boldsymbol{x^R}}, \overline{\boldsymbol{x^H}}, \boldsymbol{b^R}, \boldsymbol{b^H}, \boldsymbol{p^R}, \boldsymbol{p^H}, \overline{\boldsymbol{\tau^R}}, \overline{\boldsymbol{\tau^H}}, \overline{\boldsymbol{f^R}}, \overline{\boldsymbol{f^H}}} \quad \sum_n^N T_n \tag{9a}$$

$$\text{s.t. } \frac{\varepsilon_n \lambda_n}{F^L}\left(1-\exp\left(\overline{x_n^R}\right)-\exp\left(\overline{x_n^H}\right)\right) \le T_n,\ \forall n, \tag{9b}$$

$$\frac{d_n^R}{c} + \exp\left(\overline{\tau_n^R}\right) + \frac{\varepsilon_n \lambda_n}{F^R}\exp\left(\overline{x_n^R}-\overline{f_n^R}\right) \le T_n,\ \forall n \in \mathcal{N}, \tag{9c}$$

$$\frac{d_n^H}{c} + \exp\left(\overline{\tau_n^H}\right) + \frac{\varepsilon_n \lambda_n}{F^H}\exp\left(\overline{x_n^H}-\overline{f_n^H}\right) \le T_n,\ \forall n \in \mathcal{N}, \tag{9d}$$

$$b_n^R B_{\max}\log_2\left(1+\frac{p_n^R P_{\max} G_n^R}{b_n^R B_{\max} N_0}\right) \ge \varepsilon_n \exp\left(\overline{x_n^R}-\overline{\tau_n^R}\right),\ \forall n \in \mathcal{N}, \tag{9e}$$

$$b_n^H B_{\max}\log_2\left(1+\frac{p_n^H P_{\max} G_n^H}{b_n^H B_{\max} N_0}\right) \ge \varepsilon_n \exp\left(\overline{x_n^H}-\overline{\tau_n^H}\right),\ \forall n \in \mathcal{N}, \tag{9f}$$

$$\frac{d_n^R}{c} + \exp\left(\overline{\tau_n^R}\right) + \frac{\varepsilon_n \lambda_n}{F^R}\exp\left(\overline{x_n^R}-\overline{f_n^R}\right) \le T_n^{handoff},\ \forall n \in \mathcal{N}, \tag{9g}$$

$$\sum_{n\in\Phi_m} \exp\left(\overline{f_n^R}\right) \le 1,\ \forall m \in \mathcal{M}, \tag{9h}$$

$$\sum_{n\in\mathcal{N}} \exp\left(\overline{f_n^H}\right) \le 1, \tag{9i}$$

$$\overline{x_n^R}, \overline{x_n^H}, b_n^R, b_n^H, p_n^R, p_n^H, \overline{\tau_n^R}, \overline{\tau_n^H}, \overline{f_n^R}, \overline{f_n^H} > 0, \tag{9j}$$

$$(6c), (6d). \tag{9k}$$

命题1：由此得到的约束集 (9c)-(9i) 均为凸集。

证明。下面分别给出约束 (9c)、(9d)、(9g)、(9h)、(9i) 以及 (9e)-(9f) 的证明。

1) 约束 (9c)、(9d)、(9g)、(9h) 和 (9i) 的左侧均为正指数函数项的求和形式，因此这些约束都是凸的。

2) 约束 (9e) 和 (9f) 具有类似的表达式，如下所示：

$$\theta y \log_2(1+\frac{\eta z}{\theta}) - q\exp(\delta-\omega) \ge 0, \tag{10}$$

其中，α、β、δ 和 ω 表示优化变量，y、z 和 q 表示常数。第一项 $e(\alpha,\beta) = \alpha y \log_2(1+\frac{\beta z}{\alpha})$ 是凹函数 $\log_2(1+\beta z)$ [40] 的透视形式。因此，我们可以认为 $e(\alpha,\beta)$ 关于变量 α 和 β 联合凹。此外，容易证明，第二项 $-q\exp(\delta-\omega)$ 关于变量 δ 和 ω 联合凹。因此，最终得到的约束集 (9e) 和 (9f) 均为凸集。

□

根据证明，变量的指数变换可以有效消除变量之间的耦合。然而，由于约束 (9b) 的存在，问题 (9) 仍然是非凸的。为了解决这个问题，我们采用逐次凸近似（SCA）方法。逐次凸近似（SCA）常用于通过一阶泰勒展开迭代地近似原始非凸问题，从而将原问题转化为一系列凸问题版本 [41], [42]。在本例中，我们首先定义 $g_n = 1-\exp(\overline{x_n^R}) - \exp(\overline{x_n^H})$，其中任意 $n \in \mathcal{N}$ 均成立。然后，我们定义 $\boldsymbol{\mathcal{A}^i} = \{\mathcal{A}_n^i | n \in \mathcal{N}\}$，其中 $\hat{}\hat{}A_n^i = \{x_n^R[i], x_n^H[i]\}$ 是第 i 次迭代中的局部点集。回想一下，任意凹函数在任意给定点 [43], [44], 处都被其一阶泰勒展开全局上界，因此我们可以得到以下不等式：

$$\begin{aligned} g_n &= 1-\exp(\overline{x_n^R})-\exp(\overline{x_n^H}) \\ &\le 1-\exp\left(\hat{x_n^R}[i]\right)-\exp\left(\hat{x_n^R}[i]\right)\left(\overline{x_n^R}-\hat{x_n^R}[i]\right) \\ &\quad -\exp\left(\hat{x_n^H}[i]\right)-\exp\left(\hat{x_n^H}[i]\right)\left(\overline{x_n^H}-\hat{x_n^H}[i]\right) = \hat{g_n}[i]. \end{aligned} \tag{11}$$

通过应用SCA方法，(9b)中的凹项被一系列由给定局部点确定的线性项所替代。因此，问题(9)被转化为一系列迭代凸问题，其中第 i个凸问题为：

$$\min_{\mathrm{T}, \overline{\boldsymbol{x^R}}, \overline{\boldsymbol{x^H}}, \boldsymbol{b^R}, \boldsymbol{b^H}, \boldsymbol{p^R}, \boldsymbol{p^H}, \overline{\boldsymbol{\tau^R}}, \overline{\boldsymbol{\tau^H}}, \overline{\boldsymbol{f^R}}, \overline{\boldsymbol{f^H}}} \quad \sum_n^N T_n[i] \tag{12a}$$

$$\text{s.t.} \quad \frac{\varepsilon_n \lambda_n}{F^L}\hat{g_n}[i] \le T_n[i], \tag{12b}$$

$$(6c), (6d), (9c)-(9j). \tag{12c}$$

SCA过程如算法1所述，其计算复杂度如下所示。在每次迭代中，计算复杂度由求解凸问题决定，而凸问题可通过内点法求解。该方法需要$\log\left(\frac{N_I}{u^0\rho}\right)$

$^{\log(\xi)}$迭代次数（牛顿步），其中 N_I 是

约束数量， u 是用于估计内点法精度的初始点， γ 是停止准则， ξ 是用于更新内点法精度 [45]的参数。问题（12）中的约束数量为 $2+M+4N$，其中 M和 N分别表示系统中路侧单元和智能网联车辆的数量。因此，算法1的计算复杂度为 $\mathcal{O}\left(i_{\max}{}^{\text{对数}\left(\frac{N_L}{u^0\rho}\right)\text{对数}(\xi)}\right)$，其中 $i_{\max}$是算法1收敛所需的最大迭代次数 [4]。

算法1 求解问题(9)的基于SCA的方法。

1: 初始化精度 ζ, $i_{\max}$, $\boldsymbol{\mathcal{A}^0}$，并设置 $i=0$。
2: 重复
3: 根据给定的局部点集 $\boldsymbol{\mathcal{A}^i}$ 以及表达式(11)，得到 $\hat{}\boldsymbol{g[i]}\hat{}$。
4: 求解 the 凸 问题 (12) and 寻找最优解集 Λ ={TTT*、$\boldsymbol{x^R}$*、$\overline{\boldsymbol{x^H}}$*、$\boldsymbol{b^R}$*、$\boldsymbol{b^H}$*、$\boldsymbol{p^R}$*、$\boldsymbol{p^H}$*、$\overline{\boldsymbol{\tau^R}}$*、$\overline{\boldsymbol{\tau^H}}$*、$\overline{\boldsymbol{f^R}}$*、$\overline{\boldsymbol{f^H}}$* }.
5: 记录第 i 次迭代的目标函数值 as $\sum_n^N T_n[i]=\phi^*$，并根据 $\boldsymbol{\mathcal{A}^{i+1}}$ 进行更新 $\hat{\boldsymbol{x^R}}[i+1]=\overline{\boldsymbol{x^R}}^*$ 和 $\hat{\boldsymbol{x^H}}[i+1]=\overline{\boldsymbol{x^H}}^*$，然后 $i=i+1$。
6: 直到 $\phi^*[i]-\phi^*[i+1]\le\zeta$，或 $i>i_{\max}$。

IV. 最大值时延最小化

本节讨论一个新的优化问题，该问题通过优化智能网联车辆之间的时延公平性，聚焦于各智能网联车辆的时延性能。事实上，在时延敏感型应用中，我们必须通过最小化所有智能网联车辆的最大时延值来考虑各自的时延性能。假设参数与求和时延最小化问题相同，该优化问题可表示如下:

$$\min_{\substack{\boldsymbol{x^R},\boldsymbol{x^H},\boldsymbol{b^R},\boldsymbol{b^H},\\ \boldsymbol{p^R},\boldsymbol{p^H},\boldsymbol{f^R},\boldsymbol{f^H}}} \quad \max_{n\in\mathcal{N}} \quad \max\ \left\{T_n^L,T_n^R,T_n^H\right\} \tag{13a}$$

$$\text{s.t.} \quad T_n^R \le T_n^{handoff}, \quad \forall n\in\mathcal{N}, \tag{13b}$$

$$\sum_{n\in\mathcal{N}} b_n^R + b_n^H \le 1, \tag{13c}$$

$$p_n^R + p_n^H \le 1,\ \forall n\in\mathcal{N}, \tag{13d}$$

$$\sum_{n\in\Phi_m} f_n^R \le 1,\ \forall m\in\mathcal{M}, \tag{13e}$$

$$\sum_{n\in\mathcal{N}} f_n^H \le 1, \tag{13f}$$

$$x_n^R, x_n^H, b_n^R, b_n^H, p_n^R, p_n^H, f_n^R, f_n^H > 0. \tag{13g}$$

问题（13）通过求解优化变量的最优解，使所有智能网联车辆的最大时延值最小化。这些变量包括路侧单元的任务分割比例，即 $\mathbf{X^R}=[x_1^R,...,x_N^R]$；高空平台站的任务分割比例，即 $\mathbf{X^H}=[x_1^H,...,x_N^H]$；智能网联车辆到路侧单元链路的带宽分配，即 $\mathbf{B^R}=[b_1^R,...,b_N^R]$；智能网联车辆到高空平台站链路的带宽分配，即 $\mathbf{B^H}=[b_1^H,...,b_N^H]$；智能网联车辆到路侧单元链路的功率分配，即 $\mathbf{P^R}=[p_1^R,...,p_N^R]$；智能网联车辆到高空平台站链路的功率分配，即 $\mathbf{P^H}=[p_1^H,...,p_N^H]$；路侧单元的计算资源分配，即 $\mathbf{F^R}=[f_1^R,...,f_N^R]$；以及高空平台站的计算资源分配，即 $\mathbf{F^H}=[f_1^H,...,f_N^H]$。此外，约束（13b）表示，在路侧单元上计算任务的时延不得超过网络切换时间。与时延求和最小化问题类似，约束（13c）–（13g）分别表示基本约束，包括带宽容量、功率阈值、路侧单元和高空平台站的计算能力，以及变量的非负条件。问题（13）显然是一个非凸问题。为了解决该问题，与问题（6）的求解过程类似，我们采用以下变换和求解过程：首先，引入辅助变量 $\mathbf{T}=[T_1,...,T_N]$，将原目标函数变换为：$\min\max_{n\in\mathcal{N}} T_n$。然后，进一步引入变量 $\mathbb{T}$ 来替代 $\max_{n\in\mathcal{N}} T_n$。接下来，为处理优化变量之间的耦合关系，引入辅助变量 τ_n^R 和 τ_n^H，分别表示智能网联车辆 $n\in\mathcal{N}$ 向路侧单元和高空平台站卸载数据的时延，并添加相应的时延约束。之后，对变量 x_n^R、x_n^H、τ_n^R、τ_n^H、f_n^R 和 f_n^H 进行指数变换。在变换后的优化问题中，本地时延约束为 $\frac{\varepsilon_n\lambda_n}{F^L}\left(1-\exp\left(\overline{x_n^R}\right)-\exp\left(\overline{x_n^H}\right)\right)\le T_n$，其中左侧项为凹函数。因此，由该约束形成的集合是非凸的。为解决这一问题，我们采用算法1中提出的SCA方法，将原优化问题转化为一系列凸子问题，并通过相应的迭代得到最终结果。

V. 仿真结果

在我们的仿真中，我们考虑一条由两个路侧单元和一个高空平台站覆盖的单向道路。高空平台站部署在平流层中，海拔高度为 20 km，水平位置位于道路中央。所有智能网联车辆都可以与高空平台站即时通信，并将数据卸载到高空平台站的服务器进行数据处理。两个路侧单元分别部署在距道路起点 80 m 和 240 m 处，每个路侧单元的覆盖范围为 160 m。由路侧单元覆盖的智能网联车辆可以与路侧单元通信，并将数据卸载到其服务器进行数据处理。当智能网联车辆驶出其当前连接的路侧单元的覆盖范围时，就会发生网络切换。随后，该智能网联车辆将进入下一个路侧单元的覆盖范围，与其建立通信并获得服务。在此仿真中，初始时有十辆智能网联车辆随机部署在道路上。在.....开始时

[4]需要注意的是，我们的地面网络系统模型可以推广到多天线系统，其中每个 RSU 配备 A 根天线，并为 N 单天线智能网联汽车提供服务。为此， N 智能网联汽车可以划分为 $\frac{N}{L}$ 组，其中 L 表示每组中的智能网联汽车数量。假设 $A>\frac{N}{L}$，每个 RSU 可以形成 $\frac{N}{L}$ 个波束，且每个波束负责为一组提供服务。考虑到为每组内的 L智能网联汽车分配了正交频率资源块，所提出的带宽分配方法可以应用于每一组。

表II：仿真参数

定义	值
ICV速度（米/秒）	30
带宽容量（兆赫）	20
噪声功率谱密度（dBm/Hz）	-174
非视距路径损耗因子	3.7
莱斯因子（dB）	10
瑞利分布	$\mathcal{CN}(0,1)$
ICV发射功率（dBm）	23
HAPS定向天线增益（dBi）	17 [46]
计算密度（CPU周期/比特）	500、1,000、1,500、2,000 或 2,500
输入数据量（Kbits）	100、300、500、700 或 900
计算能力 智能网联汽车、路侧单元、高空平台站（CPU周期/秒）	2 G，32 G，100 G
最大值迭代	50
收敛精度	10^{-6}

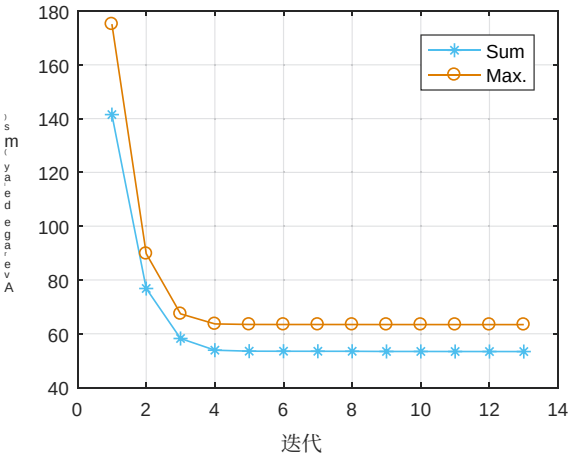


图4：采用所提方案的 SCA方法的收敛性。

每个决策时隙中，每辆智能网联车辆都会根据各自的输入数据随机生成一个待计算任务。除非另有说明，本仿真中的主要参数均按照表II进行设置。

图4展示了算法1在所提方案中求解总时延最小化问题（标记为“求和”）和最大时延最小化问题（标记为“最大值”）时的收敛性能，其中，“求和”问题的时延性能以所有智能网联车辆的平均时延衡量，而“最大值”问题则以所有智能网联车辆的最大时延值衡量。可以看出，两种时延指标的性能均在十次迭代内收敛至稳定点，从而验证了逐次凸近似方法的收敛性和效率。

图5(a)、(b)和(c)分别展示了在连续五个时间槽内，以解决“求和”问题和“最大值”问题为前提时，所提方案中智能网联车辆的平均时延、最大延迟和公平性指数。在本研究中，公平性采用广泛使用的Jain公式进行衡量，其定义为 $\mathcal{J}=(\sum_{n=1}^{N} y_n)^2/(N\times\sum_{n=1}^{N} y_n^2)$。从图中可以看出，“求和”问题的平均时延优于“最大值”问题，但其最大时延值则较差。

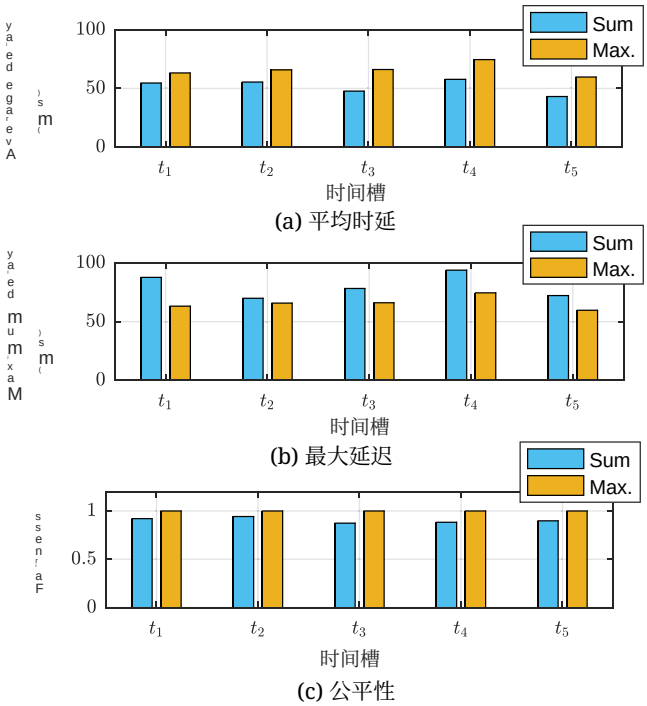


图5：所提方案的平均时延、最大延迟和公平性。

在“最大值”问题下，智能网联车辆可以获得更高的公平性值。这是因为“求和”问题的目标是优化智能网联车辆的平均时延性能，而“最大值”问题的目的是优化智能网联车辆的最大延迟。

图6–图9展示了带宽容量、智能网联汽车的发射功率阈值、高空平台站计算能力和路侧单元计算能力设置对不同方案在求解“求和”和“最大值”问题时平均时延性能的影响。为区分这些方案，我们使用“HRVIN”表示高空平台站-RSU-车辆集成网络分布式计算框架下所提方案，其中每辆智能网联汽车的任务可由其车载设备、路侧单元服务器和高空平台站服务器并行处理

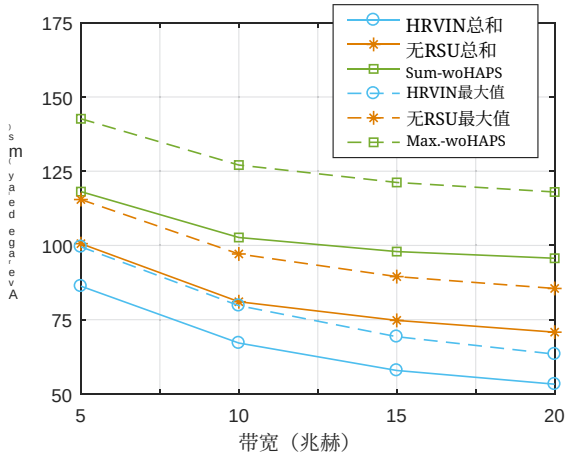


图6：平均时延与带宽的关系。

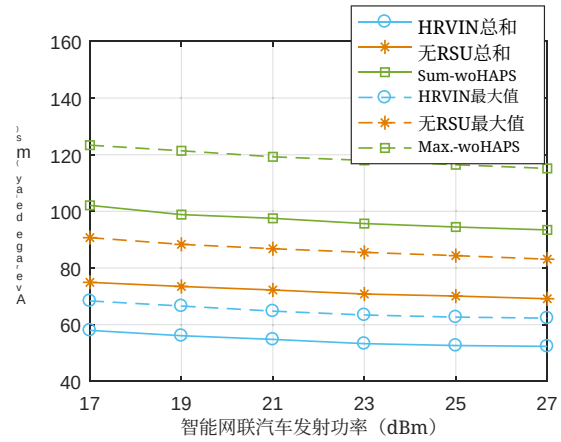


图 7：智能网联汽车的平均时延与发射功率之间的关系。

。我们使用“woRSU”表示不进行路侧单元计算的双层计算方案，并使用“woHAPS”表示不进行高空平台站计算的双层计算方案，该方案也可视为传统分布式计算方案。

图6展示了不同带宽容量设置对时延性能的影响。可以看出，与仅有两层计算资源的另外两种基准方案相比，拥有三层计算资源的“HRVIN”方案性能更佳。这表明“HRVIN”方案更有利于加速任务计算。此外，通过比较“woRSU”和“woHAPS”方案可以看出，只有采用高空平台站辅助计算的“woRSU”方案能够有效提升时延性能，且效果优于“woHAPS”方案，这是因为高空平台站能够配置比路侧单元更多的计算资源。此外，带宽容量增加能够降低计算任务的时延。这是因为带宽增加使智能网联车辆能够提升其卸载数据的能力。

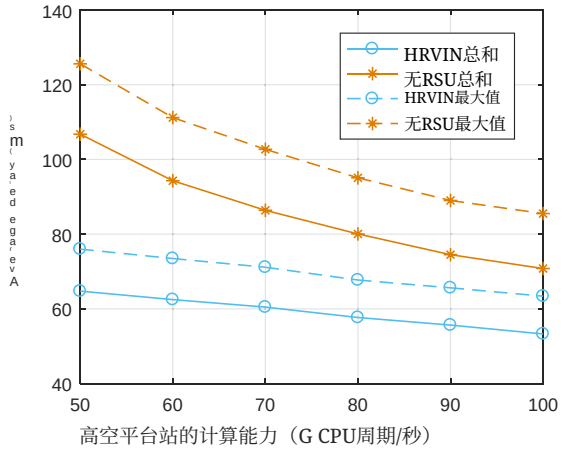


图 8：平均时延与高空平台站计算能力的关系。

到路侧单元和高空平台站。得益于此，可以卸载更多数据来加快数据处理，从而降低智能网联车辆感知到的时延。

图 7 展示了不同发射功率阈值设置对时延性能的影响。可以看出，提高功率阈值能够降低时延。与带宽容量变化的情况类似，提高功率阈值可以增强智能网联汽车卸载数据的能力。这使智能网联车辆能够卸载更多数据进行处理，从而改善时延性能。

图 8 展示了高空平台站服务器的不同计算能力设置对时延性能的影响。可以观察到，随着计算能力提高，时延逐渐降低。这表明，提高高空平台站服务器的计算能力，有助于智能网联车辆将更多数据卸载至高空平台站，以获取更强大的计算资源，从而加快数据处理。

图 9 展示了路侧单元服务器的不同计算能力设置对时延性能的影响。总体而言，提高服务器的计算能力可以加快任务的执行。

如上所述，智能网联汽车与路侧单元之间的网络切换会对计算卸载产生负面影响，而在高速移动场景下，这一问题会更加严重。为说明网络切换和移动性对计算的影响，图10(a)和(b)分别展示了在移除网络切换约束 $T_n^R \leq T_n^{handoff} \forall n \in \mathcal{N}$、求解总时延优化问题和最大时延优化问题时，不同速度设置下智能网联车辆的失败工作负载总量。我们统计平均每分钟的失败工作负载，其中工作负载定义为输入数据比特 ε (比特)与计算密度 λ (CPU周期/比特)的乘积。如图所示，总体而言，智能网联车辆的失败工作负载总量随速度增加而增加。根据公式 $T_n^{handoff} = \frac{D-|l_n \mod D|}{v}$，切换时间与智能网联汽车的速度成反比。这意味着速度越快，触发网络切换的时间就越早，因此

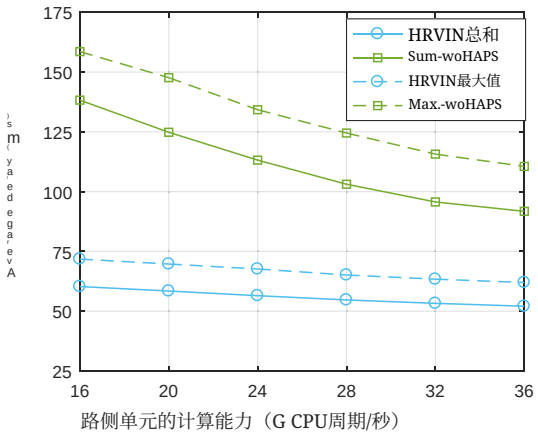


图 9：平均时延与路侧单元计算能力的关系。

行驶过程中发生切换的频率也会增加，累积的失败工作负载随之增加。图10表明，在设计计算策略时考虑网络切换因素，可以有效避免切换中断造成的故障。这对于高效、成功地进行数据处理显然意义重大，尤其是在高速移动场景下。此外，通过比较“woHAPS”和“HRVIN”方案可以看出，“woHAPS”方案的失败工作负载高于“HRVIN”方案。原因在于，不考虑切换因素时，“woHAPS”方案中的大多数智能网联车辆主要依赖路侧单元计算，因此切换造成的失败工作负载超过了“HRVIN”方案；后者可同时依赖路侧单元计算和高空平台站计算。尽管在“woHAPS”方案中考虑网络切换（即在路侧单元处加入切换约束）可以避免计算中断，但这会迫使遭遇切换的智能网联汽车只能依赖自身的本地计算，从而导致较差的时延性能。这一点将在图11中得到验证。相比之下，不考虑网络切换因素的“HRVIN”方案在一定程度上减轻了路侧单元服务器的计算负担。因此，发生切换时，被中断的工作负载更少。此外，如果“HRVIN”方案在遭遇切换时能够考虑切换因素，就可以将大量数据卸载到高空平台站，与只能依赖本地计算的“woHAPS”方案相比，可实现更低的时延。总之，这避免了网络切换对计算造成的负面影响。此外，高空平台站计算还能在避免网络切换不利影响的同时，显著提升任务计算的效率。

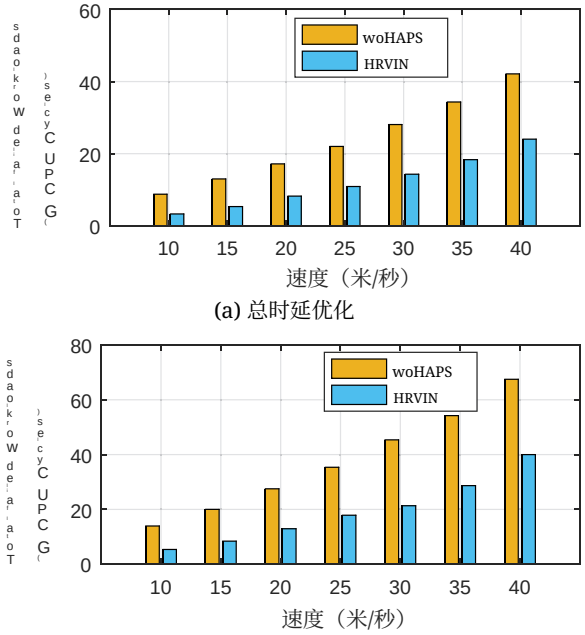


图10：因切换导致的失败工作负载总数与智能网联汽车速度的关系。

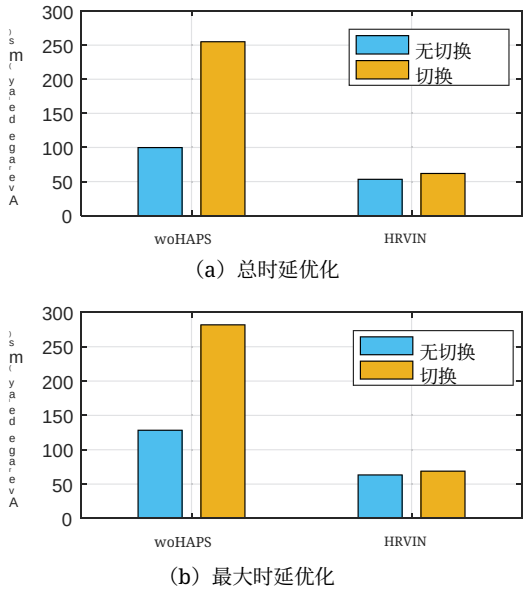


图 11：切换对时延性能的影响。

图 11 展示了两种情况下的平均时延：一种是不发生切换的情况，另一种是发生切换的情况，分别用“NoHandoff”和“Handoff”表示。在两个子图中，我们比较了“woHAPS”和“HRVIN”方案的时延性能，以说明高空平台站计算对于切换场景的重要性。正如我们所看到的，在每个子图中，边缘计算方案“woHAPS”

与“NoHandoff”情况相比，“Handoff”情况下的时延增加了近 150 ms，而“HRVIN”方案的时延增加不到 10 ms。这是因为，为避免网络切换导致计算中断，“woHAPS”方案中的智能网联汽车只能向边缘卸载少量数据，大部分数据仍需在本地计算，因此时延显著增加。然而，在“HRVIN”方案中，智能网联汽车可以将路侧单元无法处理的数据发送至

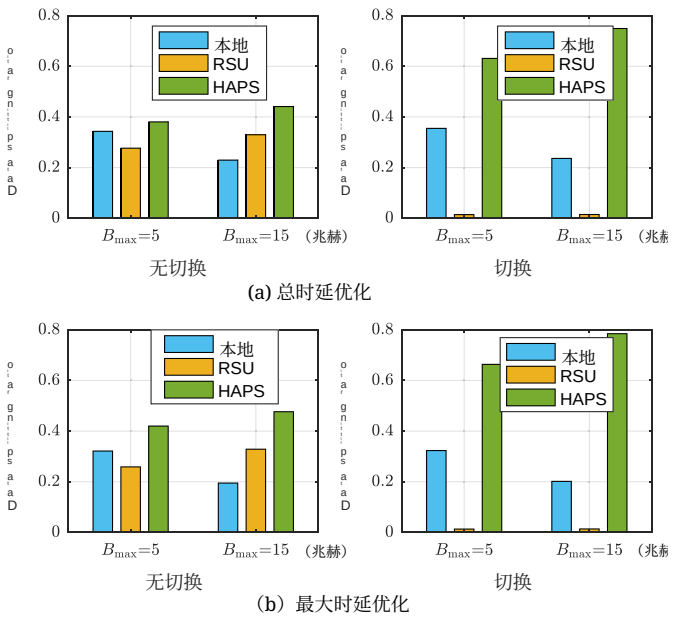


图12：无切换和切换情况下，数据分割比例随带宽的变化。

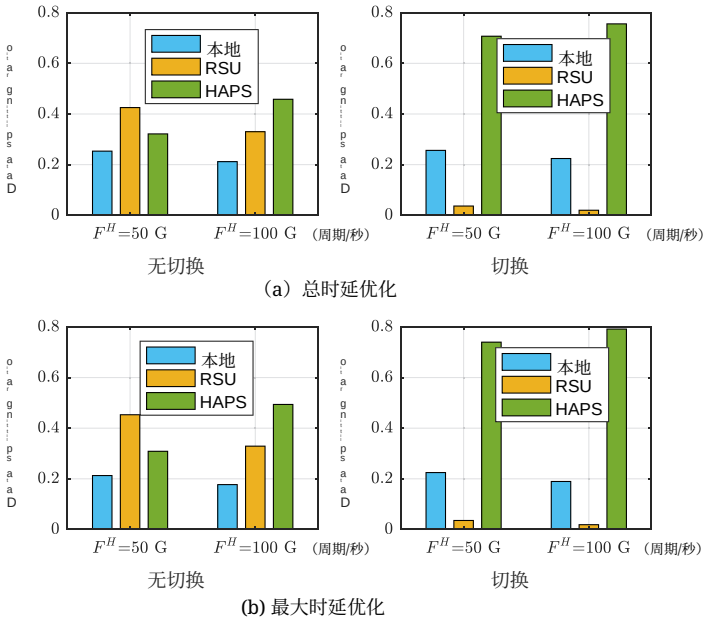


图 13：无切换和切换情况下的数据分割比例与高空平台站服务器计算能力的关系。

高空平台站，因此时延增幅较小且可以接受。通过上述对比可以看出，传统边缘计算在网络切换情况下的时延性能会突然恶化，这显然不利于智能网联汽车行驶的稳定性和安全性。相比之下，高空平台站可以通过在可接受的时延增幅内完成任务计算，消除网络切换带来的负面影响。因此，引入高空平台站计算有助于智能网联车辆应对网络切换。

图12展示了所提方案在不同带宽设置下，任务在三个计算节点（本地、路侧单元和高空平台站）上的平均处理比例。图12（a）和（b）展示了在求解总时延优化问题的前提下，“无切换”和“切换”两种情况下数据拆分的对比。图12（c）和（d）反映了求解最大时延优化问题时的上述情况。从子图（a）和（b）中的“无切换”情况可以看出，当带宽为5兆赫时，数据大多由本地或高空平台站服务器处理。这是因为较小的带宽限制了智能网联车辆卸载数据的能力。因此，数据倾向于由智能网联汽车的车载设备直接处理，或卸载至计算资源更充足的高空平台站。随着带宽容量增至15兆赫，向路侧单元和高空平台站卸载数据的吞吐量显著提升，有效降低了相应的通信时延，因此本地计算的划分比例下降，而路侧单元计算和高空平台站计算的划分比例上升。此外，对比“无切换”和“切换”两种情况可以发现，路侧单元计算的划分比例从25%–35%显著降至3%。这是因为卸载至路侧单元的数据必须在网络切换触发前成功处理。

因此，在“切换”情况下，路侧单元的数据分割比例较低。与此同时，高空平台站的划分比例从38%–46%显著升至62%–78%，这表明路侧单元的大部分工作负载已转移至高空平台站，以避免计算中断。这说明，发生切换时，高空平台站计算发挥着重要作用。

类似地，图13展示了在所提方案中，不同高空平台站计算能力设置下，任务在三个计算节点（本地、路侧单元和高空平台站）上执行的平均比例。如子图（a）和（c）中的“无切换”情形所示，当高空平台站的计算能力设为50 G CPU周期/秒时，路侧单元计算发挥着至关重要的作用。这是因为高空平台站的计算能力低于系统中两个路侧单元的总计算能力（每个路侧单元均为 32 G CPU周期/秒）。当高空平台站服务器的计算能力增加到 100 G CPU周期/秒时，可以明显看出，高空平台站计算发挥了重要作用。此外，通过比较“无切换”和“切换”两种情形可以发现，发生切换时，高空平台站承担了大部分工作负载（约为 70%～80%），而路侧单元计算仅处理一小部分工作负载。综合图12和图13，可以得出两个结论：第一，在“切换”情形下，大部分数据已从路侧单元迁移到高空平台站，因此高空平台站在此情形下发挥着至关重要的作用，这证实了该分布式框架中的高空平台站计算能够有效应对网络切换。第二，根据数据划分比例，所提分布式计算方案确实能够实现更灵活、自适应的任务调度。

VI. 结论

在本研究中，我们提出了一种借助高空平台站的分布式并行计算方案，以降低车辆计算任务的时延，同时避免切换带来的负面影响，为车辆提供平稳、稳定的服务体验。正如我们所见，该方案将数据灵活地划分为三部分，并由智能网联车辆、路侧单元和高空平台站并行处理。通过将路侧单元计算的任务部分设定为在网络切换前完成，该方案有效消除了切换对计算造成的负面影响。在此基础上，本研究建立了总时延优化问题，并采用SCA方法求解。随后，我们还讨论了最大时延优化问题的建模与求解。最后，大量仿真结果证实了所提方案的有效性。

高空平台站网络的能耗问题至关重要，因为悬停和计算都需要消耗能量。考虑到高空平台站的能耗问题，其在计算方面的作用将减弱，从而延长智能网联汽车感知到的时延。未来，我们将探讨这一全面而有趣的课题。

参考文献

[1] J. Zhou、D. Tian、Y. Wang、Z. Sheng、X. Duan 和 V. C. Le- ung，“网联车辆系统中的可靠性最优协作通信与计算”，IEEE Transactions on Mobile Computing，第 19 卷，第 1216–1232 页，2020 年 5 月。[2] K. N. Qureshi、S. Din、G. Jeon 和 F. Piccialli，“车联网：关键技术、网络模型、解决方案、挑战与未来展望”，IEEE Transactions on Intelligent Transportation Systems，第 22 卷，第 1777–1786 页，2021 年 3 月。[3] M. Chen、Z. Yang、W. Saad、C. Yin、H. V. Poor 和 S. Cui，“面向无线网络联邦学习的联合学习与通信框架”，IEEE Transactions on Wireless Communications，第 20 卷，第 269–283 页，2021 年 1 月。[4] Z. Ning、P. Dong、X. Wang、M. S. Obaidat、X. Hu、L. Guo、Y. Guo、J. Huang、B. Hu 和 Y. Li，“深度强化学习遇上 5G车联网：面向交通大数据的分布式卸载框架”，IEEE Transactions on Industrial I nformatics，第 16 卷，第 1352–1361 页，2020 年 2 月。[5] M. Chen、D. G¨ und¨uz、K. Huang、W. Saad、M. Bennis、A. V. Feljan 和 H. V. Poor，“无线网络中的分布式学习：最新进展与未来挑战”，IEEE Journal on Selected Areas in Communications，第 39 卷，第 3579–3605 页，2021 年 12 月。[6] J. Kang、X. Li、J. Nie、Y. Liu、M. Xu、Z. Xiong、D. Niyato 和 Q. Yan，“面向物联网人工智能的通信高效、跨链赋能联邦学习”，IEEE Transactions on Network Science and Engineering，第 9 卷，第 2966–2977 页，2022 年 5 月。[7] Y. Ye、R. Q. Hu、G. Lu 和 L. Shi，“利用上行非正交多址接入提升移动边缘计算网络中的时延约束计算性能”，IEEE Transactions on Communications，第 68 卷，第 2409–2425 页，2020 年 4 月。[8] M. Chen、N. Shlezinger、H. V. Poor 和 S. Cui，“通信高效联邦学习”，美国国家科学院院刊，第 118 卷，2021 年 4 月。[9]NTT DATA 白皮书，“当汽车接管驾驶：一窥自动驾驶的未来”。访问日期：2022 年 6 月 10 日。[在线]。网址：
https://us.nttdata.co
m/en/-/media/assets/white-paper/mfg-autonomous-car
s-white-paper.pdf.

[10] F. Lyu、H. Zhu、N. Cheng、H. Zhou、W. Xu、M. Li 和 X. Shen，“面向可靠安全应用的城市车对车通信特性分析”，IEEE Transactions on I ntelligent Transportation Systems，第 21 卷，第 2586–2602 页，2019 年 6 月。[11] G. Qiao、S. Leng、K. Zhang 和 Y. He，“车载边缘多接入网络中的协同任务卸载”，IEEE Commun. Mag.，第 56 卷，第 48–54 页，2018 年 8 月。

[12] W. Shi、J. Li、W. Xu 和 H. Zhou，“边缘计算：愿景与挑战”，《IEEE 物联网杂志》，第 32 卷，第 130–137 页，2018 年 6 月。[13] M. LiWang、S. Dai、Z. Gao、X. Du、M. Guizani 和 H. Dai，“5G 卫星—地面车联网架构下具有时延和成本约束的计算卸载激励机制”，《IEEE 无线通信》，第 26 卷，第 124–132 页，2019 年 8 月。[14] H. Du、D. Niyato、Y.-A. Xie、Y. Cheng、J. Kang 和 D. I. Kim，“干扰器辅助多天线无人机隐蔽通信的性能分析与优化”，IEEE Journal on Selected Areas in Communications，第 40 卷，第 2962–2979 页，2022 年 10 月。[15] Z. Jia、M. Sheng、J. Li 和 Z. Han，“面向 6G 空天地一体化网络中的数据采集与传输：高空平台站与低轨卫星协作方案”，《IEEE 物联网期刊》，第 9 卷，第 10516–10528 页，2022 年 7 月。[16] Z. Jia、Q. Wu、C. Dong、C. Yuen 和 Z. Han，“通过高空平台站与无人机协作实现物联网分层空中计算”，《IEEE 物联网期刊》，2022 年，提前访问。[17] ITU，“无线电规则条款。”
http://www.itu.int/pub/R-REG-RR-2016，2016 年。[18] Z. Jia、M. Sheng、J. Li、D. Zhou 和 Z. Han，“6G 中的高空平台站接入与低轨卫星回传联合设计：基于匹配博弈的方法”，IEEE Journal on Selected Areas in Communications，第 39 卷，第 1147–1159 页，2021 年 4 月。[19] W. Jaafar 和 H. Yanikomeroglu，“HAPS-ITS：为跨洲高速公路提供未来智能交通系统服务”，《IEEE 通信杂志》，第 60 卷，第 80–86 页，2022 年 10 月。[20] “空中风能的前景与挑战。”访问日期：2021 年 5 月 18 日。[在线]。网址：
htt
ps://physicsworld.com/a/the-promise-and-challenges-of-airborne-wind-energy/

[21]“HAPSMobile 与 APB 就使用全聚合物电池为高空平台站开发储能电池达成基本协议。”访问日期：2021年5月18日。[在线]。网址：
https://www.hapsmobile.com/en/news/press/2020/20201224 01/. [22] G. K. Kurt、M. G. Khoshkholgh、S. Alfattani、A. Ibrahim、T. S. J. Darwish、M. S. Alam、H. Yanikomeroglu 和 A. Yongacoglu，“面向未来的高空平台站（高空平台站）网络：愿景与框架”，IEEE Communications Surveys & Tutorials，第23卷，第729–779页，2021年第二季度。[23] Z. Jia、Q. Wu、C. Dong、C. Yuen 和 Z. Han，“通过高空平台站与无人机协作实现物联网分层空中计算”，IEEE Internet of Things Journal，2022年。（提前出版）。[24] G. K. Kurt 和 H. Yanikomeroglu，“面向下一代空中配送网络的通信、计算、缓存与感知：以高空平台站作为赋能技术”，IEEE Vehicular Technology Magazine，第16卷，第108–117页，2021年9月。[25] Q. Ren、O. Abbasi、G. K. Kurt、H. Yanikomeroglu 和 J. Chen，“高空平台站辅助智能交通系统中的缓存与计算卸载”，IEEE Transactions on Wireless Communications，第21卷，第9010–9024页，2022年11月。[26] S. Yu、X. Gong、Q. Shi、X. Wang 和 X. Chen，“EC-SAGIN：边缘计算增强的空天地一体化车联网”，IEEE Internet of Things Journal，第9卷，第5742–5754页，2022年4月。[27] V. B. Souza、M. H. Pereira、L. H. S. Lelis 和 X. Masip-Bruin，“通过智能跨域切换提升车载雾计算中的资源可用性”，载于 GLOBECOM 2020——2020 IEEE 全球通信大会，第1–6页，2020年12月。[28] T. M. Ho 和 K.-K. Nguyen，“多接入边缘计算无线网络中的联合服务器选择、协同卸载与切换：一种深度强化学习方法”，IEEE Transactions on Mobile Computing，第21卷，第2421–2435页，2022年7月。[29] Q. Yuan、J. Li、H. Zhou、T. Lin、G. Luo 和 X. Shen，“面向车载边缘计算的联合服务迁移与移动性优化方法”，IEEE Transactions on Vehicular Technology，第69卷，第9041–9052页，2020年8月。[30] W. Zhan、C. Luo、J. Wang、C. Wang、G. Min、H. Duan 和 Q. Zhu，“基于深度强化学习的车载边缘计算卸载调度”，IEEE Internet of Things Journal，第7卷，第5449–5465页，2020年6月。[31] 张H、R. Wang、W. Sun 和 H. Zhao，“区块链超密集边缘计算的移动性管理：一种深度强化学习方法”，IEEE Transactions on Wireless Communications，第20卷，第7346–7359页，2021年11月。[32] T. Ojima 和 T. Fujii，“利用用户移动性预测进行移动边缘计算资源管理”，载于2018年国际信息网络会议（ICOIN），第718–720页，2018年1月。

[33] C. W. Lee、L. M. Chen、M. C. Chen 和 Y. S. Sun，“基于移动 IPv6 的无线覆盖网络切换框架”，IEEE Journal on Selected Areas in Communications，第23卷，第2118–2128页，2005年11月。[34] M. Liu、Z. C. Li 和 X. B. Guo，“WWAN 与 WLAN 之间垂直切换的高效决策算法”，《计算机科学与技术杂志》，第22卷，第114–120页，2007年1月。[35] K. Shafiee、A. Attar 和 V. C. M. Leung，“车载异构网络中的最优分布式垂直切换策略”，IEEE Journal on Selected Areas in Communications，第29卷，第534–544页，2011年3月。[36] A. Ibrahim 和 A. S. Alfa，“利用拉格朗日松弛法进行高空平台无线资源分配”，IEEE Transactions on Wireless Communications，第14卷，第5823–5835页，2015年6月。[37] J. Kang、H. Du、Z. Li、Z. Xiong、S. Ma、D. Niyato 和 Y. Li，“面向任务的语义通信中的个性化显著性：图像传输与性能分析”，IEEE Journal on Selected Areas in Communications，第41卷，第186–201页，2023年11月。[38] A. Alsharoa 和 M. S. Alouini，“通过资源优化的卫星—机载—地面一体化网络提升全球连通性”，IEEE Transactions on Wireless Communications，第19卷，第5088–5100页，2020年4月。[39] S. Karapantazis 和 F. Pavlidou，“通过高空平台实现宽带通信：综述”，IEEE Communications Surveys & Tutorials，第7卷，第2–31页，2005年第一季度。[40] S. Boyd 和 L. Vandenberghe，《凸优化》。剑桥大学出版社。[41] J. Kaleva、A. T¨olli 和 M. Juntti，“通过逐次凸近似实现具有 QoS 约束的干扰广播信道分布式和速率最大化”，IEEE Transactions on Signal Processing，第64卷，第2788–2802页，2016年6月。[42] A. Liu、V. K. N. Lau 和 M. Zhao，“两阶段随机非凸优化的在线逐次凸近似”，IEEE Transactions on Signal Processing，第66卷，第5941–5955页，2018年11月。[43] Q. Ren、J. Chen、O. Abbasi、G. K. Kurt、H. Yanikomeroglu 和 F. R. Yu，“一种应用驱动的非正交多址接入计算卸载方案”，IEEE Internet of Things Journal，第8卷，第1453–1466页，2021年2月。[44] Q. Wu、Y. Zeng 和 R. Zhang，“多无人机辅助无线网络的联合轨迹与通信设计”，IEEE Transactions on Wireless Communications，第17卷，第2109–2121页，2018年3月。[45] O. Abbasi、H. Yanikomeroglu、A. Ebrahimi 和 N. M. Yamchi，“具有动态 NOMA/OMA 的无人机辅助 NR-V2X 网络轨迹设计与功率分配”，IEEE Transactions on Wireless Communications，第19卷，第7153–7168页，2020年11月。[46] F. B. Mismar 和 B. L. Evans，“由 6 GHz 以下 LTE 信令辅助的 mmWave 新空口部分盲切换”，载于2018年 IEEE 国际通信会议研讨会（ICC Workshops），第1–5页，2018年5月。