



能量信息深度融合的移动网络架构及其技术挑战

周盛^{1*}, 孙宇璇², 姜之源³, 龚杰⁴, 常征⁵, 牛志升¹

1. 清华大学电子工程系, 北京 100084

2. 北京交通大学电子信息工程学院, 北京 100044

3. 上海大学通信与信息工程学院, 上海 200444

4. 中山大学计算机学院, 广州 510275

5. 电子科技大学计算机科学与工程学院, 成都 611731

* 通信作者. E-mail: sheng.zhou@tsinghua.edu.cn

收稿日期: 2023-10-08; 修回日期: 2023-12-30; 接受日期: 2024-02-22; 网络出版日期: 2024-04-11

国家自然科学基金 (批准号: 62341108, 62301024, 62271300, 62171481, 62071105) 资助项目

摘要 5G 的大规模部署伴随着网络能耗呈指数级增长, 与此同时, 国家提出了“碳达峰”与“碳中和”的“双碳”目标, 因此移动网络的节能减排已刻不容缓. 然而, 一方面随着人工智能等技术的发展, 未来 6G 网络中计算类业务预期将迎来大幅增长; 另一方面, 实现绿色低碳需要提升可再生能源的利用率, 但绿色新能源供能不稳难以保证服务质量. 为应对以上挑战, 本文提出一种能量与信息服务深度融合的低碳网络架构, 通过构建一个协调异质资源的智能控制面, 充分利用并调度移动网络中泛在的算力、通信和储能, 在保证 6G 个性化业务服务质量要求的同时, 以期大幅降低网络能耗与碳排放. 本文同时还讨论了该架构下的调度机制设计以及可能的未来研究方向.

关键词 6G, 低碳, 节能, 通信计算融合, 绿色能源

1 引言

随着 5G 加快部署, 移动网络中基站部署密度越来越大, 且由于大规模天线等空口技术的引入, 5G 单站的能耗是 4G 的 2~4 倍^[1], 导致网络能耗呈指数级增长, 节能减排已经刻不容缓. 另一方面, 近年来国家制定了“建设资源节约型和环境友好型社会”以及“碳达峰”、“碳中和”的“双碳”目标. 因此, 迫切需要新的网络架构, 在提高移动通信系统容量的同时, 有望大幅降低全网能耗和碳排放.

目前, 学术界和工业界提出了多种新型空口技术来降低传输能耗. 例如, 超大规模多输入多输出 (multiple-input multiple-output, MIMO) 技术通过更窄的发射波束, 在终端接收相同功率的情况下, 基站可以配置更低的发射功率. 有测算表明, 在边缘用户覆盖体验相同的前提下, 超大规模 MIMO 技术

引用格式: 周盛, 孙宇璇, 姜之源, 等. 能量信息深度融合的移动网络架构及其技术挑战. 中国科学: 信息科学, 2024, 54: 745–757, doi: 10.1360/SSI-2023-0290
Zhou S, Sun Y X, Jiang Z Y, et al. Deep integration of the energy and information: network architecture and technology challenges (in Chinese). Sci Sin Inform, 2024, 54: 745–757, doi: 10.1360/SSI-2023-0290

可以降低 50% 以上的发射能耗^[2]。又例如, 智能射频通道选择技术被用来减少使用的射频通道数, 从而降低射频功耗^[3]。然而, 这些新型空口技术具有较高的计算复杂度, 在降低传输能耗的同时增加了处理能耗。同时, 有研究表明, 即使小区中没有任何业务需求, 由于基站仍需要处于工作状态以保证网络的基础覆盖 (即保证用户可以随时随地接入网络的信号覆盖, 从而消耗大量的射频和基带能耗, 用于广播信令并处理用户接入请求), 基站也会消耗大量的能量。据中国电信的测算, 现有蜂窝网络基础覆盖的能耗约占全网实际能耗的 50% 以上, 因此仅仅降低传输能耗是远远不够的, 需要在架构上实现革新。

低能耗蜂窝网络的架构革新, 源于学界和业界对基站动态休眠的探索。蜂窝网络的基站部署是为了满足峰值业务需求, 因此当业务需求较低时, 可通过适时的基站休眠节约网络能耗。基于基站动态休眠的节能机制, 可根据不同基站的负荷信息、空口能力、活跃的终端数量及移动性、小区利用率等, 判断是否将基站/小区休眠或开启^[4~6]。然而在传统蜂窝网络架构中, 一个基站兼备网络覆盖、用户接入控制以及数据传输等多重身份, 引入低负载基站的休眠固然可以降低网络能耗, 但也会影响网络的覆盖和服务质量。为此, 第一个针对蜂窝网络节能的超蜂窝 (hyper cellular) 网络新架构被学者提出^[7], 其核心是控制信令覆盖与业务数据覆盖的适度解耦, 即永远在线的控制基站保障信令覆盖并协调业务基站, 在支持低负载业务基站按需动态休眠的同时最大化活跃基站的服务效率, 从而降低全网能耗, 实现能量效率与频谱效率的联合优化^[8]。进一步地, 超蜂窝架构与 C-RAN 云架构无线接入网概念相结合, 采用集中的计算资源池实现高效的空口信号处理, 利用软件定义实现异质资源重构, 可演进成为软件定义超蜂窝架构^[9]。同时注意到, 5G 网络的 CU/DU 分离, 与超蜂窝架构的控制与业务分离, 具有相似的概念。由此可见, 无论在核心网 (对应于软件定义网络) 还是接入网, 控制与业务服务分离都是架构演进的趋势。

此外, 预计 6G 将承载大量的虚拟现实 (virtual reality, VR)、增强现实 (augmented reality, AR)、分布式人工智能 (artificial intelligence, AI) 等计算密集型业务^[10], 同时, 3GPP 在 2020 年启动了 NG-RAN 的人工智能及机器学习研究, 重点针对 5G 网络节电、负载均衡和移动性优化这 3 个应用场景, 分析引入人工智能的可行性及潜在解决方案, 这些都包括了边缘计算的支持。因此基于边缘计算的低碳设计受到越来越多的关注, 人们致力于解决计算效率的最大化和设备能耗的最小化问题^[11]。例如, 学者们研究联合任务卸载和计算资源分配的移动边缘计算系统的能量效率最大化问题^[12]。然而, 目前对通信调度和计算任务卸载的处理是相对分立的, 一方面难以解决空口信号的实时处理, 另一方面也无法通过协同通信和计算实现通信、计算、能量等异质资源的精细化管理, 获得更高的节能增益。以上两点对网络的架构也提出了新要求。

更重要的是, 为了切实达到“双碳”目标, 仅仅依靠蜂窝网络能耗的“节流”显然是不够的, 势必要实现能量的“开源”, 即有效利用可再生能源等绿色能源, 与电网协同对蜂窝网供电, 以期大幅降低全网的碳排放。然而, 可再生能源在时空上存在动态性、能量到达随机变化、能量分布不均, 因此难以保证可靠的服务。为此, 近年来学者们研究了动态能量供给下的无线通信理论与方法, 挖掘能量状态信息 (即能量到达统计特性和电池状态) 对无线资源管理、传输调度与用户服务的作用机理^[13, 14]。全球各大运营商也开始部署可再生能源供电的基站。然而, 目前可再生能源的利用率仍较低, 其根本原因是储能与供电无法实现与信息服务在时空维度细颗粒度的融合与适配。近年来, 随着自动驾驶和车联网的高速发展, 智能网联汽车将具备强大的环境感知、计算和通信能力; 同时, 大力发展新能源汽车是推动我国交通运输绿色低碳转型的重要举措。这些新能源汽车将为网络注入海量的数据和算力^[15, 16], 并作为储能节点提供移动的绿色能源。因此, 急需实现移动环境中能量与信息的双向互动, 通过利用泛在且移动的计算和能量资源获得节能增益。为此, 本文试图充分利用移动智能在 6G 智能化、绿色

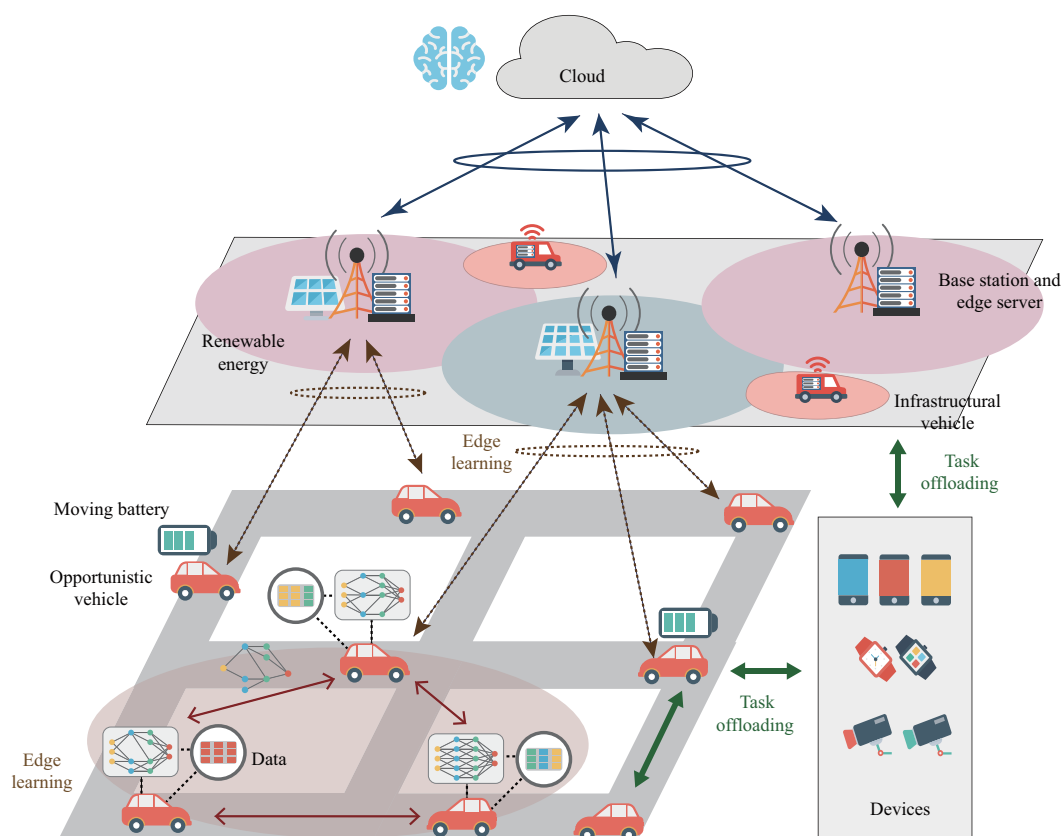


图 1 (网络版彩图) 能量与信息服务深度融合的网络架构

Figure 1 (Color online) Network architecture that integrates the energy and information services

化进程中发挥的积极作用, 提出一种能量与信息服务深度融合的低碳移动网络架构, 该架构的创新性要素包括低碳异构移动网络智能控制面、泛在的算力、通信和储能能力以及通信计算一体化的领域定制硬件架构. 同时, 本文初步探索了该架构支撑下移动性增强的通信-计算-能量资源智能调度策略. 最终针对这一新的架构, 讨论可能的低碳节能增益与技术挑战.

2 能量与信息服务深度融合的低碳移动网络架构

本文所提出的低碳网络架构, 力图支撑电网、可再生能源, 及新能源汽车等移动储能节点协同供电, 如图 1 所示. 进一步提出对应的服务机制, 使其能够高效利用时空高度动态的异质能量, 满足多样化业务的通信和计算服务需求. 面向 6G 典型智能业务, 基于多种通信链路协同 (车车、车路等) 和分布式计算等技术, 设计多智能体协作、车网协同的新型服务机制, 以高效利用移动智能的泛在算力、通信和储能能力. 量化可再生能源到达模型和储能节点的移动性模型, 分析能量在时间和空间上的动态分布特性, 定量表述能量的不确定性; 另一方面, 分析移动环境中的通信链路质量和信息交互融合增益. 在此基础上, 从能量、信息融合的角度, 分析能量和信息双向互换的两方面约束条件, 研究能量不确定性条件下的网络能量信息协同机理.

为了实现动态能量和信息服务的深度融合, 支撑电网与绿色能源的协作, 协同调度各代蜂窝网、WIFI、可见光通信和有线网等多种网络和接入方法, 充分发挥各自在能效上的优势, 需要构建适应性和兼容

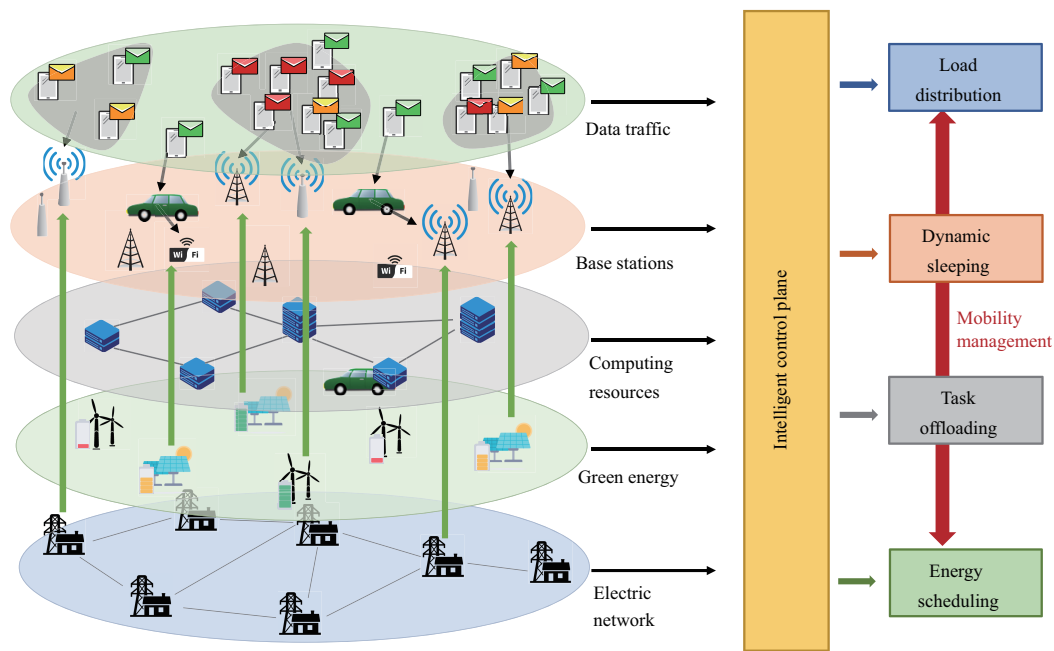


图 2 (网络版彩图) 低碳异构移动网络智能控制面概要

Figure 2 (Color online) Intelligent control plane for low carbon heterogeneous network

性强的智能控制面, 感知并优化网络资源管理和动态休眠, 降低网络运营和部署阶段的碳排放. 该架构的主要特点和要素详述如下.

2.1 低碳异构移动网络智能控制面

本文作者前期提出的超蜂窝网络可以保障蜂窝网络的无缝覆盖和频谱效率, 同时, 通过引入基站的动态休眠和资源调度, 可大幅度降低网络整体能耗. 在超蜂窝网络架构中, 控制基站承担了网络控制面的角色, 但控制基站只感知网络的负载变化, 尚无法感知网络的碳排放状态, 这在引入可再生能源之后显得尤为关键. 同时, 随着计算资源和人工智能算法在网络边缘的大量引入, 移动网络将面临更加严峻的管理成本与能耗挑战. 因此, 亟须设计具备能耗碳排放感知能力的智能蜂窝网络控制面 (如图 2), 决定能耗碳排放感知的“探针”形态和更新频率; 基于储能数字化技术和能量流的量化表征, 在网络控制面统一调配异质能量, 适配通信资源和信息服务; 运用网络功能虚拟化的思路, 虚拟化无线和计算资源, 使移动网络根据实际业务需求进行动态重构, 支持灵活的基站/载波休眠关断、网络重构、剪裁及网络升级等, 与核心网的 SDN 控制器取得良好的协调, 提升频谱和能量的利用率; 设计统一的控制面信令, 兼容各代蜂窝网、WIFI、可见光通信和有线网等多种接入网的协同操作.

以降低碳排放为目标的移动网络, 在低能耗网络运行的同时, 必然也关注部署阶段的成本与碳排放, 尽量延长已有设备的使用时间, 这对网络的兼容性和可重构、可升级特性提出了要求. 针对网络中存在尚在服役“多代同堂”的 2G/3G/4G/5G 基站, 设计合理方式刻画其服务能力、覆盖能力和能量效率, 基于此协同优化这些基站与 WIFI、可见光通信和有线网等多种接入方法的网络部署、动态资源管理和节点休眠等机制, 适配异质资源, 以最小的部署代价实现低能耗的网络运行, 并指导可再生能源和储能单元的部署优化. 此外, 考虑 6G 可能采纳的新空口技术在低碳网络部署和运行中的角色与性能增益, 解决 5G 基站部署过于密集的问题. 例如, 智能反射面不需要进行信号处理而是只进行电

磁波反射, 因此其能耗仅存在于控制电路, 是一种较为低功耗的网络覆盖补充方案, 相比于部署中继, 具有相对更低的通信时延和更简单的硬件设备需求。

2.2 泛在的算力、通信和储能能力

在自动驾驶、车联网、新能源汽车等领域快速发展的背景下, 新一代汽车将具备环境感知、计算、通信、自主决策能力和电池储能, 成为无线网络边缘的超级移动智能终端。一方面, 行驶或停泊的车辆, 特别是公共汽车、出租车等公共交通工具, 可共享其车载计算平台的剩余算力, 通过 LTE V2X 或 5G V2X 等车联网通信协议接收网络中的计算任务并提供机会式的计算服务, 因而我们将这些车辆称为“机会车”(opportunistic vehicle)。同时, 机会车可通过环境感知实时收集与自动驾驶、交通态势及无线通信等应用相关的数据, 并通过分布式机器学习、联邦学习、多智能体强化学习等方式将海量数据及时转化为信息和知识, 形成边缘智能。机会车的绿色能源供电也有益于网络的节能减排。同时, 充分挖掘和利用机会车在时间、空间的富余计算、存储、通信资源, 将大大提升社会数字基础设施的能力和效率。

另一方面, 运营商也可在部署传统静态基础设施的基础上, 协同部署专用车辆作为基站或边缘服务器, 称为“基础设施车”(infrastructural vehicle)。由于业务量往往随时间和空间维起伏变化, 无线网络无需同时满足全域的峰值流量需求。通过静态节点和专用车的协同部署, 可降低基站和边缘服务器的总数量, 并缓解基站站址紧缺问题; 在可再生能源供电基站发生能量中断时, 通过专用车的储能携能, 可保证网络覆盖和服务能力。为了充分发挥移动智能的节能减碳能力, 需要对业务量、可再生能源到达率、机会车算力等进行智能预测, 并在此基础上协同优化专用车轨迹、位置和基站动态休眠策略, 在满足业务服务需求的同时, 最小化网络整体能耗和总碳排放。

2.3 通信计算一体化的领域定制硬件架构

未来 6G 移动通信网络中, 智能计算将成为不可或缺的组成要件。移动边缘计算和无线通信将深度融合, 以提供泛在的智能。一方面, 移动网络将提供给用户多种智能服务, 例如边缘智能处理和联邦学习等, 这要求移动网络具备处理低延时边缘智能计算的能力; 另一方面, 网络的运行也将基于内生智能, 通过自组织、自优化、自演进来不断提高网络的运行效率, 而且绿色低碳网络的运行大多依赖于智能算法的应用, 进一步提高了智能计算在移动网络中的重要性。综上所述, 未来移动通信需要通过有限的硬件资源和能耗, 在支持日益增长的无线基带信号处理需求之外, 支撑强大的智能计算能力, 因此亟须解决如何有效提升能量和资源效率, 协同优化通信和计算。

本文针对图计算、机器学习、无线通信基带处理等应用中数据量与计算量大、计算需求快速迭代且多样化的问题, 提出基于 RISC-V 指令集扩展的领域专用处理器与多核芯片架构。在微架构层面, 设计与标量核紧耦合的向量扩展计算阵列、高并发向量重组单元, 以提高单模块的数据级并行与数据交互带宽; 在多核系统层面, 提出支持分级调度机制的多层次并行架构, 通过 L1 级调度器与多计算簇并行以支撑线程级数据流处理, 通过 L2 级调度器与多向量阵列并行以支撑程序块级数据流处理。基于此种架构, 无线通信的信号处理和智能计算领域中的多种应用均可以共享一套领域专用硬件架构, 大大提升了资源和能量的利用效率, 有利于实现真正低碳的移动网络运行。具体架构图以及与其他主流架构的对比如图 3 所示。

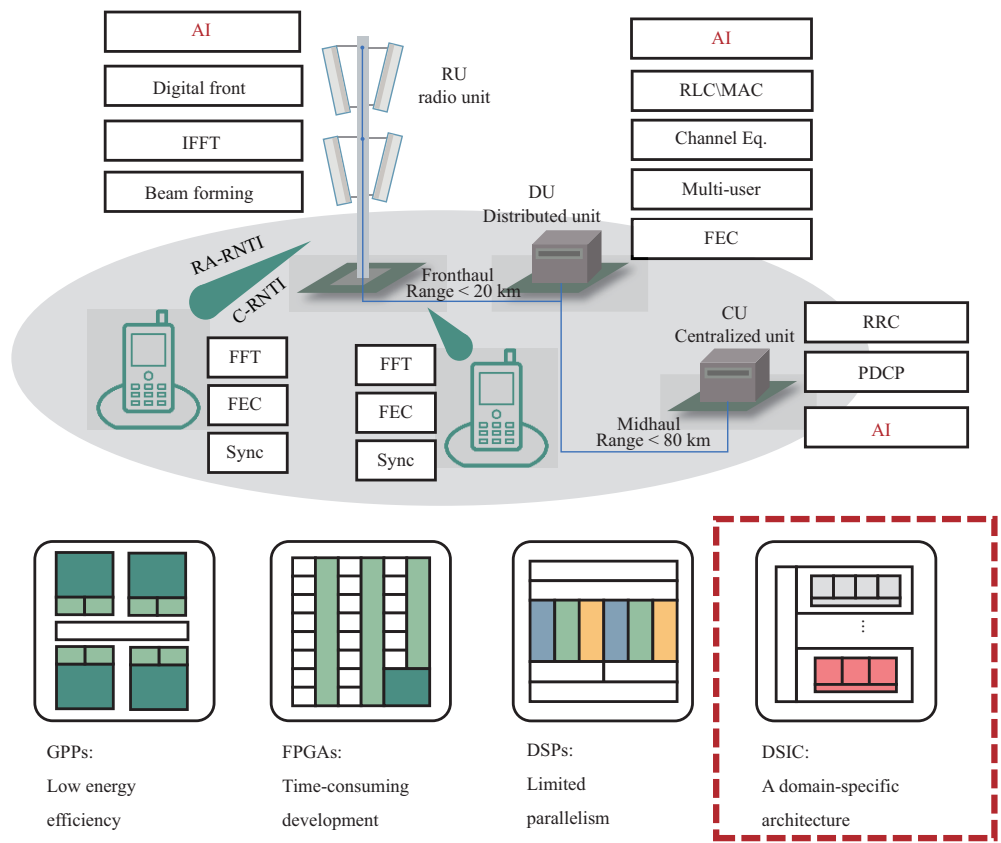


图 3 (网络版彩图) 通信计算一体化的领域专用硬件架构
Figure 3 (Color online) Domain specific hardware architecture for unified communication and computing

3 移动性增强的通信 – 计算 – 能量资源智简调度

本节将基于上述网络架构, 探讨如何高效利用移动智能的泛在算力、通信和储能能力, 通过精准调度动态受限的通信、计算、绿色能源等异质资源实现低能耗、低碳的 6G 智能服务, 揭示移动性增强的绿色信号处理与业务服务机理. 重点关注能量、信息的融合交互机理及其与移动性的耦合关联, 并在此基础上探索通信 – 计算 – 能量资源的智简调度和移动性管理. 具体包括以下 4 个方面的研究内容与挑战.

3.1 移动条件下的能量信息融合交互机理

区别于以往以频谱效率、延时等为指标的业务需求模型, AR/VR、实时监控与推断、分布式机器学习等 6G 智能业务模型需考虑更多因素, 如图 4 所示. 实现网络节能减排的前提是探明能量与信息的融合交互机理, 建立业务需求与能耗、碳排放的映射关系. 在业务需求方面, 车联网、物联网中的机器间通信业务, 需要在延时的基础上, 进一步考虑信息新鲜度、信息时效性需求等, 由此引入信息年龄 (age of information)、信息紧迫度 (urgency of information)^[17] 等性能指标. AR/VR 等实时会话类业务, 要考虑传输精度、推理精度与实时性之间的折中. 对于分布式机器学习任务, 要考虑在线推理的实时性需求以及分布式训练的精度、收敛性需求等, 还需综合考虑传输与计算的效率和资源利用率. 同时, 智能业务处理涉及中央处理器 (central processing unit, CPU)、图形处理器 (graphics processing unit,

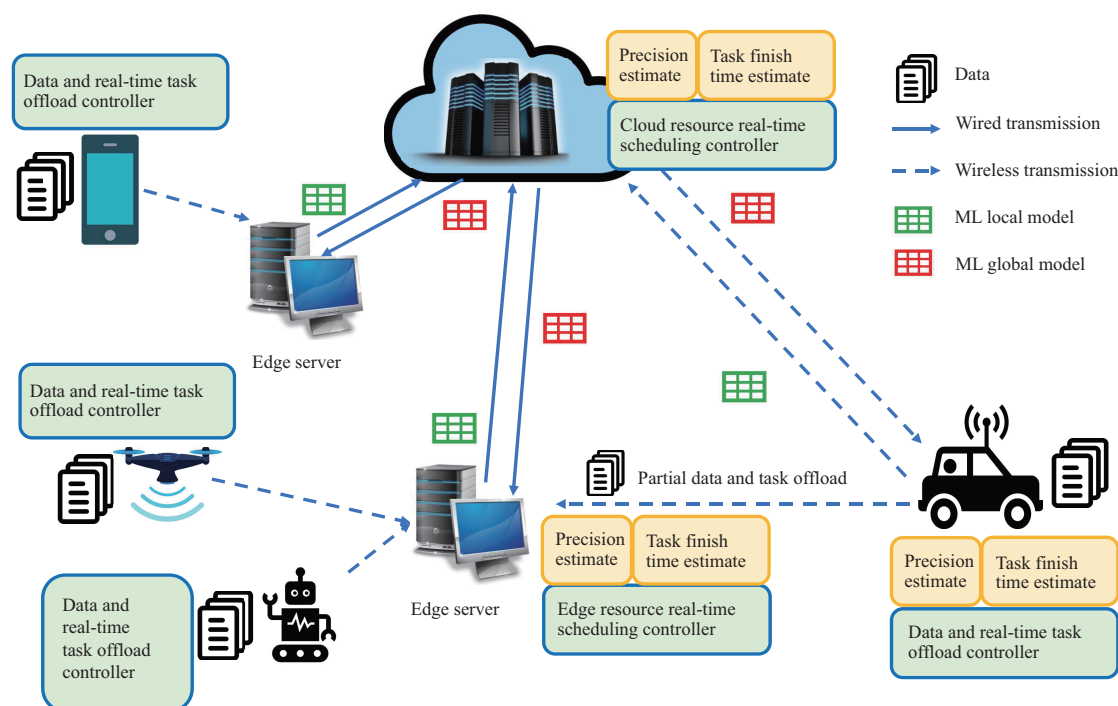


图4 (网络版彩图) 6G 智能业务需求及服务流程示意图

Figure 4 (Color online) Illustration of the requirements and service flows of 6G intelligent traffic

GPU) 及定制硬件等异构计算平台, 需要建立不同业务在不同平台的能耗和碳排放模型, 对业务的性能和资源需求进行量化建模.

智能业务的实时性、精度等性能需求往往随情境变化而改变, 且移动网络中各节点的通信计算资源和能量状态是动态时变的. 为此, 需要设计灵活智简的智能业务服务机制, 遵循“按需适度服务”原则, 改变任务的通信量、计算量和能耗, 使业务服务能够自适应于多维时变的资源和性能需求并节能减排. 对于推理任务, 基于神经网络剪枝、动态神经网络技术, 依次对推理任务的整体计算复杂度及中间层传输数据量进行压缩, 以支持灵活的网络分割, 优化终端和网络侧能效. 根据无线接入端计算资源的分布、计算处理的时延、信道状态及实时性需求, 选择任务卸载的最优分割点和卸载策略. 对于多任务场景, 通过设计负载聚集和 GPU 批处理机制可获得更多的休眠机会. 对于训练类任务, 考虑终端计算能力和传输性能的双重随机性对分布式机器学习收敛速度及模型精度的影响, 基于此设计并优化终端的训练迭代次数、部分参数更新策略, 以及基于量化、稀疏化等方法的轻量化梯度传输调度与带宽分配. 本方向上的前代表性工作包括文献 [18~22] 等.

在此基础上, 研究智能业务与移动性的动态适配理论与方法, 基于移动性和智能业务需求优化任务拆分和轻量化程度, 使推理、训练等典型任务的资源需求能够适配于多个移动智能体的动态拓扑和异构资源约束, 以满足任务精度和实时性需求, 同时降低能耗. 具体地, 给定智能体拓扑关系和移动性模型, 设计动态组网与路由机制, 提出适配于移动性以及各智能体数据质量和资源约束的任务拆分方法. 进一步, 利用深度学习对移动性进行预测, 并基于此对任务的执行流程和拆分进行时序优化, 最大化任务成功完成概率.

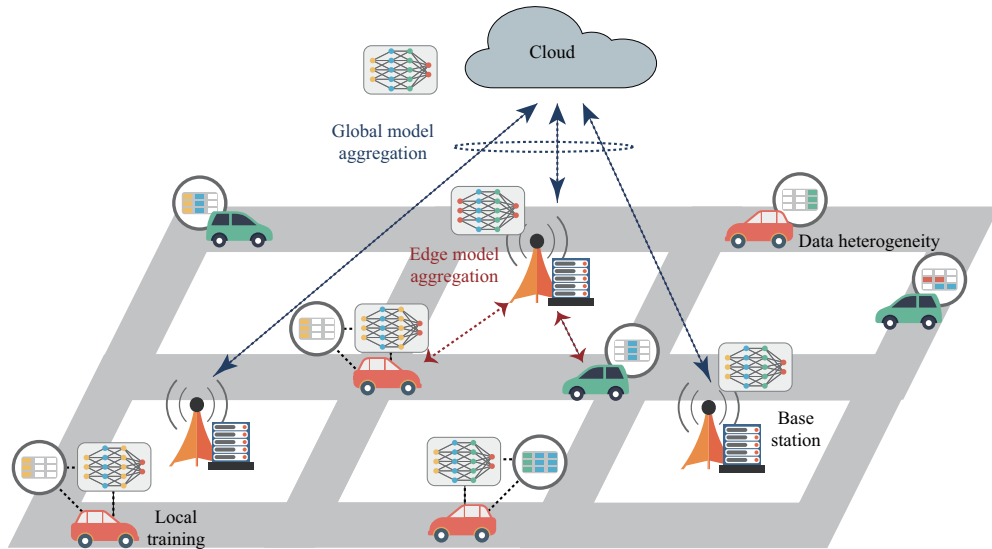


图 5 (网络版彩图) 云-边-车分层联邦学习系统

Figure 5 (Color online) Cloud-edge-vehicle hierarchical federated learning system

3.2 移动性增强的业务服务机理

所提网络架构中的移动性增益主要来自于以下两方面.

(1) 更多通信和计算机会: 机会车可以通过 V2X Sidelink 与其他车辆或终端进行通信, 赋能端对端计算任务卸载和分布式机器学习. 机会车也可作为中继, 扩展网络覆盖范围、提高传输速率. 由于移动性可带来更近的通信距离、更好的信道条件和视距通信机会, 若能抓住高通信速率的传输时机, 可同时降低传输功率和延时. 另一方面, 尽管移动性减少了机会车与固定位置基站或移动终端之间的连接时间, 但在相对速度接近的情况下, 车辆间的拓扑反而变化较慢, 因而可以探索车辆间的分布式、协同计算机制, 提高业务服务质量和能效.

(2) 促进数据融合: 在分布式学习和推理任务中, 数据质量直接影响了任务的难度和精度. 在分布式学习中, 受到数据收集时间、位置等因素影响, 车辆本地数据和全局数据分布可能不同, 即具有数据异质性 (data heterogeneity), 影响训练精度^[23]. 随着移动性的增加, 具有不同分布和特征数据会扩散到全网, 每个车辆也将与携带其他分布数据的车辆相遇并交互信息, 由此通过数据融合降低数据异质度, 加快分布式学习的收敛速率, 因而降低训练能耗. 目标检测和跟踪等推理任务可通过多视角数据融合提高精度、降低复杂度. 例如, 在自动驾驶中, 相邻车辆可通过共享感知数据进行协同感知, 有效解决单车感知中的遮挡问题. 在移动条件下, 邻车更有可能提供性能增益较大的视角, 提高目标检测精度, 或在精度不变时降低检测算法复杂度, 从而节能减排.

在前期代表性工作^[16, 24, 25]中, 我们通过数值仿真初步验证了移动性增益. 具体地, 在文献[24]中, 考虑车辆间的分布式计算任务卸载, 所提出的基于移动性和业务量的协同计算策略, 可大幅降低计算任务的完成延时, 提高具有延时界约束的计算任务的成功完成概率. 在文献[16]中, 考虑移动车辆在基站的协同下, 通过联邦学习优化智能波束选择策略, 基于 Raymobtime 数据集¹⁾的仿真实验表明, 中等移动性 (15 m/s) 的训练收敛速率最快 (优于 7.5 m/s 和 25 m/s). 文献[26]进一步考虑了车辆的能量状态, 设计了通信、计算和能量三维约束下的移动性感知联邦学习策略.

1) <https://www.lasse.ufpa.br/raymobtime/>.

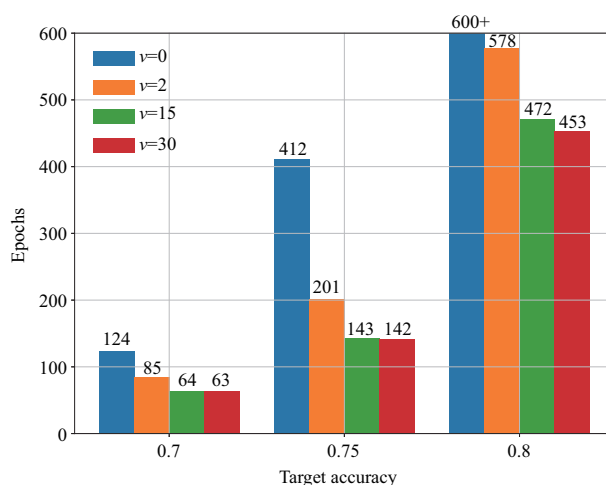


图6 (网络版彩图) 在不同车速下达到各目标训练精度的训练轮数

Figure 6 (Color online) Number of training epochs to reach the target accuracy under different speeds

在文献 [25] 中, 考虑如图 5 所示的云-边-车(端) 分层联邦学习, 车辆的初始数据分布具有区域性异质特征, 即每个边缘下所有车辆具有相同数据, 不同边缘的车辆数据异质. 同时, 在此结构当中, 边缘节点和车的能量状况会对联邦学习的性能产生影响. 边缘节点需要对模型进行聚合, 而在端侧需要进行模型训练, 都需要消耗能量, 所以边和端的能量状况都会对联邦学习的表现有一定影响. 当对联邦学习终端进行选择时, 就需要对终端能量进行考虑, 选择合适的终端来保证系统性能. 在训练过程中, 车辆在云覆盖范围内匀速运动并跨越各个边缘的覆盖范围, 造成异质数据的融合. 在训练总时长受限的条件下, 基于 CIFAR-10 数据集进行仿真实验. 如图 6 所示, 仿真结果表明移动性可明显提升模型的训练速度, 降低到达各个目标训练精度的训练轮数. 特别地, 当目标精度为 75% 时, 车辆静止时需要 412 轮训练达到该精度, 车辆速度为 1 m/s 时需要 201 轮训练, 降低了 51.2%; 车辆速度为 30 m/s 时仅需 142 轮训练, 相比静止时降低了 65.5%.

未来需通过理论分析揭示移动性对于服务性能提升和全网节能的积极影响. 例如, 分析移动环境中的通信链路质量和信息交互融合增益, 分布式机器学习任务中数据异质度随移动性的变化及其对训练收敛速率和模型精度的影响. 进而与能量信息交互机理相结合, 分析车辆、边缘服务器的能量状态对调度策略和学习性能的影响, 进而探明移动性的节能减排增益.

3.3 通信-计算-能量资源智简调度

在移动条件下, 通信链路质量动态时变, 各智能体在满足自身高优先级业务后所剩余的可共享算力也是动态时变的, 导致通信速率、算力、能量等状态信息获取开销大、准确度低. 该挑战的一个解决思路是设计低复杂度的在线学习算法, 对动态时变的移动智能体资源状态进行在线学习, 从而简化信令开销和资源调度流程, 实现异质资源的智简调度. 例如, 在前期代表性工作 [24] 中, 我们考虑车辆间的分布式任务卸载, 基于强化学习中的多臂赌博机 (multi-armed bandit, MAB) 理论设计了卸载算法. 算法的核心是通过观察计算任务卸载至不同服务车节点的总卸载延时, 学习不同服务车节点的延时性能均值或分布; 通过引入服务车节点出现时间的变量自适应于动态网络拓扑; 通过对所学习的延时分布进行离散化, 降低算法复杂度. 基于真实道路环境的仿真结果显示, 在给定卸载延时界为 0.5 s, 任务到达率和服务率之比为 0.2 时, 卸载可靠性可达 99.6%. 在此基础上, 未来工作需综合考虑延时、时效

性等业务性能需求和能耗、碳排放等指标, 设计在线学习的回报函数, 并针对移动性和动态拓扑, 对学习算法开展针对性设计与优化, 使其在动态环境中快速收敛, 且复杂度低、运行速度快.

进一步, 任务完成精度、可靠性等性能与情境紧耦合, 如联合感知精度与智能体遮挡关系、位置、速度等多维因素有关, 分布式学习模型准确度与数据分布及异质程度紧密相关, 难以通过理论公式闭式表征. 为此, 需研究基于在线学习的情境感知与任务性能表征方法. 在此基础上, 融合凸优化、随机优化等理论与方法, 设计低复杂度的通信 - 计算 - 能量协同调度与优化, 在满足业务需求的同时, 降低总能耗. 在前期代表性工作^[27]中, 面向自动驾驶联合感知任务, 我们提出了基于无休止多臂赌博机 (restless MAB) 理论的在线学习算法, 对可综合反映传输速率、观测视角和数据质量的联合感知增益进行在线学习; 基于网联自动驾驶数据集 DOLPHINS^[28]的仿真结果显示, 在给定延时界为 0.1 s、协同感知车辆比例较高时, 所提算法可将协同检测增益提高 10%, 目标检测精度提高 3.1%.

3.4 基于能量信息双向互动的智简移动性管理

智能体移动会导致执行计算任务的虚拟机迁移, 产生额外的资源重构时延. 传统基于信道质量的切换方式会导致频繁的虚拟机迁移, 将不再适用. 针对网络在拓扑、能量获取、频谱信道、计算资源等方面的多维动态特性, 亟须提出智简移动性管理方法, 基于移动性和资源预测, 设计高动态环境下的计算任务匹配迁移与高能效切换. 在前期代表性工作^[29]中, 我们引入了计算卸载的切换代价, 即当任务上传和结果下载由不同节点服务时, 会产生额外的延时开销. 而后, 我们联合考虑了车辆移动性和通信计算能力, 提出延时最小化的卸载策略.

进一步, 通过移动性规划, 对任务时效、能效及交通效率开展联合规划, 在保证交通效率需求的同时, 对能量、算力、数据进行按需空间搬移, 提升通信链路稳定性及质量, 通过能量与信息的双向互动, 进一步提升任务完成性能并降低能耗. 例如, 绿色能源利用率更高、电量更充足的车辆可优化轨迹, 主动靠近发出任务卸载请求的终端或基站; 具有数据融合需求的车辆可维持相对稳定的拓扑和较近距离等; 车辆还能够对可再生能源供电基站进行能量和算力补充, 降低可再生能源供电网络中的能量中断概率.

4 挑战与未来研究方向

能量信息深度融合网络架构的实现, 还存在很多挑战. 我们从能耗模型、接入机制、传输架构、计算资源配置等方面, 提出实现上述网络架构的未来研究方向.

(1) 蜂窝网络系统级能耗模型与能量、信息、计算资源的融合交互机理. 建立基于器件的设备级能耗模型, 结合分析 6G 候选技术的关键算法、核心器件和系统架构等因素, 对基站组成器件提炼关键能耗影响参数, 建立器件级的能耗回归和缩放模型, 搭建系统级能耗模型. 同时, 考虑网络承载业务的多样性, 对业务流依照其执行流程与服务质量需求进行划分, 建立业务需求与网络能耗及总功率开销的对应关系. 按照业务的计算需求, 将其划分为数据通信业务与计算处理业务; 按照业务的延时需求, 将其划分为实时业务与非实时业务. 通过建立网络承载业务流的数学模型, 分析不同业务流混合情况对于能量供应的差异化需求. 相比于传统电能, 可再生能源供应具有不确定性, 对网络服务质量将产生不利影响. 通过对可再生能源的随机产生过程进行随机性建模, 分析不同类型业务的服务质量边界, 探明能量随机情况下的网络服务能力极限, 刻画实时业务与非实时业务之间的服务优先级关系. 考虑涉及多次通信交互与计算处理环节的业务, 分析各环节能量开销并进行资源分配, 建立针对一般业务模型的服务能力边界分析框架.

(2) 绿色能源感知的大规模免授权随机接入调度优化. 建立低时延高可靠免授权随机接入的能耗模型, 包括基于超大规模天线阵列的射频能耗、基于免授权随机接入的基带处理能耗, 以及用户侧接入与传输能耗. 基于这一模型揭示射频能耗和处理能耗的关系, 为设计低时延高可靠的免授权随机接入机制提供理论依据. 进一步, 通过分布式 MIMO 的部署为设备接入提供良好的信噪比条件, 使用少量的射频链路实现射频能耗大幅降低, 利用轻量化 AI 从少量训练数据中学习先验知识, 设计低复杂度的轻量化 AI 检测算法显著减少基带处理能耗. 在免授权接入机制中, 探明导频长度、导频复用、动态资源分配等对传输能耗和基带处理能耗的影响, 并基于此设计绿色能源感知的免授权灵巧接入方法. 最后, 针对不同的网络通信场景, 定义不同网络情境下的模型特征和参数, 考虑终端间直接通信模式协同辅助传统的终端-基站通信模式, 以降低能耗. 终端直连模式可同时节省基站的处理、传输能耗以及终端的接入能耗, 是异构场景下高能效的接入方式, 为大规模可靠接入提供了可行性保障.

(3) 具有服务质量保障的大规模低能耗传输架构. 大量实测数据表明, 基站能耗占整个网络能耗的 80% 以上. 基站能耗, 除了固有的配套设备和机房空调能耗外, 主要由传输能耗和处理能耗组成, 其中传输能耗主要包括射频能耗, 由射频链路能耗与数量决定, 而处理能耗指基带算法的计算能耗, 由浮点运算的数量决定. 随着超大规模多输入多输出天线阵列的使用, 基站的传输能耗和处理能耗随之增加. 需探明大规模空口传输和处理算法的能耗关系, 构建服务质量保障条件下的传输和处理能耗的折中关系. 研究基于分布式 MIMO 的边缘计算技术, 探明多点协作式资源分配方案, 利用小样本机器学习技术设计以元强化学习为基础的计算卸载策略, 实现通信系统传输能耗的下降和资源利用率的提高; 研究分布式 MIMO 的高智能传输技术, 通过对业务量的时间、空间动态变化的感知以及对用户通信需求的智能预测, 利用低复杂度的智能算法, 实现分布式天线单元的合理调度, 在低开销低能耗情况下, 探究广覆盖的多点协作式混合预编码算法.

(4) 绿色能源的时空变化对边缘节点资源分配的作用机理. 考虑可再生能源收集的随机性、不确定性, 分析边缘设施的电力与可再生能源的协同供电模式, 设计能量采集及存储设备与计算单元的联合部署方案. 类似于计算资源, 能量的获取对应也有云(中心式)与边缘(分布式)两种. 考虑将云边结合的思想应用于异质资源的联合部署, 研究能量收集设备的空间部署和时间到达的内在特性, 与算力网络异构资源的配置结合, 并根据网络实际的地理、成本、需求等约束条件, 合理选择纯绿色能源供电模式与电网绿色能源混合供电模式, 设计最优的异质资源部署方案. 在给定的网络部署下, 考虑节点间通过电网传递剩余能量或负载转移, 实现能量的等效共享, 提高可再生能源利用率. 结合节点能量状态和资源状况选择合适的卸载策略, 通过本地计算和边缘卸载等策略的选择, 优化系统内任务执行所造成的边缘服务器端电网电力能源成本及云服务器端云资源租用成本, 研究能量、传输、计算以及存储等异质资源的智能协同调度方法. 例如, 在任务卸载中, 不论是终端还是任务节点都需要对能量情况进行考虑. 如果终端的能量不足以支撑任务的执行, 就需要进行任务卸载. 类似地, 如果任务节点的能量不够, 终端就需要判断是否卸载以及如何选择卸载的目的地. 在考虑任务卸载决策之时, 有必要将承载任务节点的能量状态作为问题条件之一, 以充分考虑能量情况而做出最优的选择策略. 任务节点当中的任务排队情况会对任务卸载执行的时延产生较大影响, 因此当任务队列较长之时, 终端就会考虑其卸载决策, 包括是否卸载或者选择其他节点进行卸载. 所以当用户数过多或者节点服务速率较低时, 需要对任务排队进行分析, 优化卸载决策. 研究能量存储与数据缓存在信息时效性约束下的深度融合及对计算任务准确度的影响, 尤其是在考虑终端移动性的基础之上, 设计基于能量域和数据域存储的异质资源动态匹配方法. 在终端位置, 从能量与数据内容的匹配程度入手, 运用匹配论等方法找到最优方式来选择推送策略和提高能量使用效率.

5 结论

本文面向国家“双碳”目标,提出了一种能量与信息服务深度融合的低碳网络架构,该架构充分利用移动网络中泛在的算力、通信和储能,在保证 6G 个性化业务服务质量要求的同时,以期大幅降低网络能耗与碳排放.本文介绍了该架构的主要元素,即异质资源智能控制面板、泛在的算力、通信和储能,以及通信计算一体化的领域定制硬件架构.本文还通过一个实例阐述了如何利用移动智能提升学习任务等新业务的服务质量,即如何设计移动性增强的通信-计算-能量资源智能调度策略以提升资源利用率,从而有助于网络能耗的降低.最后,本文讨论了所提架构相关的研究挑战与未来研究方向.

参考文献

- 1 I C L, Han S F, Bian S. Energy-efficient 5G for a greener future. *Nat Electron*, 2020, 3: 182–184
- 2 Prasad K N R S V, Hossain E, Bhargava V K. Energy efficiency in massive MIMO-Based 5G networks: opportunities and challenges. *IEEE Wireless Commun*, 2017, 24: 86–94
- 3 van Nguyen M S, Do D T, Al-Rubaye S, et al. Exploiting impacts of antenna selection and energy harvesting for massive network connectivity. *IEEE Trans Commun*, 2021, 69: 7587–7602
- 4 Zhou S, Gong J, Yang Z, et al. Green Mobile Access Network With Dynamic Base Station Energy Saving. *IEICE, Technical Report*, IA2009-50, 2009
- 5 Gong J, Zhou S, Niu Z S. A dynamic programming approach for base station sleeping in cellular networks. *IEICE Trans Commun*, 2012, E95-B: 551–562
- 6 Guo X Y, Niu Z S, Zhou S, et al. Delay-constrained energy-optimal base station sleeping control. *IEEE J Sel Areas Commun*, 2016, 34: 1073–1085
- 7 Niu Z S, Zhou S, Zhong S D, et al. Energy efficiency and resource optimized hyper-cellular mobile communication system architecture and its technical challenges. *Sci Sin Inform*, 2012, 42: 1191–1203 [牛志升, 周盛, 周世东, 等. 能效与资源优化的超蜂窝移动通信系统新架构及其技术挑战. *中国科学: 信息科学*, 2012, 42: 1191–1203]
- 8 Zhou S, Zhang S, Niu Z S. Energy efficient network planning and dynamic control for hyper-cellular network. *Sci Sin Inform*, 2017, 47: 1566–1582 [周盛, 张珊, 牛志升. 超蜂窝网络的高能效基站部署及其动态控制. *中国科学: 信息科学*, 2017, 47: 1566–1582]
- 9 Zhou S, Zhao T, Niu Z S, et al. Software-defined hyper-cellular architecture for green and elastic wireless access. *IEEE Commun Mag*, 2016, 54: 12–19
- 10 Saad W, Bennis M, Chen M. A vision of 6G wireless systems: applications, trends, technologies, and open research problems. *IEEE Network*, 2020, 34: 134–142
- 11 Ning Z L, Kong X J, Xia F, et al. Green and sustainable cloud of things: enabling collaborative edge computing. *IEEE Commun Mag*, 2019, 57: 72–78
- 12 Zhao T C, Zhou S, Sun Y X, et al. A predictive frame transmission scheme for cloud gaming in mobile edge cloudlet systems. *IEEE Trans Mobile Comput*, 2023, 22: 3774–3789
- 13 Gong J, Zhou S, Niu Z S. Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems. *IEEE Trans Commun*, 2013, 61: 3040–3049
- 14 Gong J, Zhou S, Zhou Z Y, et al. Policy optimization for content push via energy harvesting small cells in heterogeneous networks. *IEEE Trans Wireless Commun*, 2017, 16: 717–729
- 15 Zhou S, Sun Y X, Jiang Z Y, et al. Exploiting moving intelligence: delay-optimized computation offloading in vehicular fog networks. *IEEE Commun Mag*, 2019, 57: 49–55
- 16 Sun Y X, Xie B W, Zhou S, et al. MEET: mobility-enhanced edge intelligence for smart and green 6G networks. *IEEE Commun Mag*, 2023, 61: 64–70
- 17 Zheng X, Zhou S, Niu Z S. Urgency of information for context-aware timely status updates in remote control systems. *IEEE Trans Wireless Commun*, 2020, 19: 7237–7250
- 18 Sun Y X, Shi W Q, Huang X F, et al. Edge learning with timeliness constraints: challenges and solutions. *IEEE Commun Mag*, 2020, 58: 27–33

- 19 Huang X F, Zhou S. Dynamic compression ratio selection for edge inference systems with hard deadlines. *IEEE Internet Things J*, 2020, 7: 8800–8810
- 20 Shi W Q, Zhou S, Niu Z S, et al. Multiuser co-inference with batch processing capable edge server. *IEEE Trans Wireless Commun*, 2023, 22: 286–300
- 21 Shi W Q, Zhou S, Niu Z S, et al. Joint device scheduling and resource allocation for latency constrained wireless federated learning. *IEEE Trans Wireless Commun*, 2021, 20: 453–467
- 22 Sun Y X, Zhou S, Niu Z S, et al. Dynamic scheduling for over-the-air federated edge learning with energy constraints. *IEEE J Sel Areas Commun*, 2022, 40: 227–242
- 23 Li X, Huang K X, Yang W H, et al. On the convergence of FedAvg on non-IID data. In: *Proceedings of the 8th International Conference on Learning Representations*, 2020
- 24 Sun Y X, Zhou S, Niu Z S. Distributed task replication for vehicular edge computing: performance analysis and learning-based algorithm. *IEEE Trans Wireless Commun*, 2021, 20: 1138–1151
- 25 Chen T, Yan J, Sun Y, et al. Data-heterogeneous hierarchical federated learning with mobility. In: *Proceedings of International Teletraffic Congress*, 2023
- 26 Zhang X R, Chang Z, Hu T, et al. Vehicle selection and resource allocation for federated learning-assisted vehicular network. *IEEE Trans Mobile Comput*, 2023. doi: 10.1109/TMC.2023.3283295
- 27 Jia Y K, Mao R Q, Sun Y X, et al. MASS: mobility-aware sensor scheduling of cooperative perception for connected automated driving. *IEEE Trans Veh Technol*, 2023. doi: 10.1109/TVT.2023.3285308
- 28 Mao R Q, Guo J Y, Jia Y K, et al. DOLPHINS: dataset for collaborative perception enabled harmonious and interconnected self-driving. In: *Proceedings of the 16th Asian Conference on Computer Vision, Macao*, 2023. 495–511
- 29 Nan Z J, Zhou S, Jia Y J, et al. Joint task offloading and resource allocation for vehicular edge computing with result feedback delay. *IEEE Trans Wireless Commun*, 2023, 22: 6547–6561

Deep integration of the energy and information: network architecture and technology challenges

Sheng ZHOU^{1*}, Yuxuan SUN², Zhiyuan JIANG³, Jie GONG⁴, Zheng CHANG⁵ & Zhisheng NIU¹

1. *Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;*

2. *School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;*

3. *School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China;*

4. *School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510275, China;*

5. *School of Computer Science and Engineering, University of Electronic Science and Technology, Chengdu 611731, China*

* Corresponding author. E-mail: sheng.zhou@tsinghua.edu.cn

Abstract The large-scale deployment of 5G is accompanied by an exponential increase in network energy consumption. At the same time, the country has put forward the “dual-carbon” strategic requirements of “carbon peak” and “carbon neutrality”, so it is urgent to achieve energy-saving and emission reduction in mobile networks. However, on the one hand, with the development of technologies such as artificial intelligence, computation-intensive services are expected to see significant growth in future 6G networks. On the other hand, achieving green and low-carbon goals requires improving the utilization of renewable energy, but the instability of green energy supply makes it difficult to guarantee service quality. To address these challenges, this paper proposes a low-carbon network architecture that deeply integrates energy and information services. By building an intelligent control plane that coordinates heterogeneous resources, it fully utilizes and schedules the ubiquitous computing, communication, and energy storage in mobile networks. The goal is to greatly reduce network energy consumption and carbon emissions while ensuring the quality of personalized services in 6G. This paper also discusses the design of scheduling mechanisms under this architecture and possible future research directions.

Keywords 6G, low carbon, energy saving, communications and computing integration, green energy