

# 面向“双碳”战略的绿色通信与网络：挑战与对策

牛志升<sup>1,2</sup>, 周盛<sup>1,2</sup>, 孙宇璇<sup>1,2</sup>

(1. 清华大学电子工程系, 北京 100084; 2. 北京信息科学与技术国家研究中心, 北京 100084)

**摘要:** 面向国家“碳达峰”与“碳中和”的“双碳”战略需求, 移动通信与网络需要在满足不断增长的业务需求前提下大幅度降低全网能耗, 因此需要研究使用更少的能量传递更多信息 (SMILE, send more information bits with less energy) 的理论与技术。为了应对该挑战, 仅靠无线传输技术的改进和硬件实现水平的提高是远远不够的, 需要从系统和网络的角度探索能量的高效利用机理与方法。从能量的“节流”和“开源”2个维度展开, 并针对日益增长的计算能耗给出解决方案。具体地, 通过引入超蜂窝网络架构实现网络的柔性覆盖与弹性接入, 使业务基站和边缘服务器在业务量较低时可以进入休眠状态, 减少能量的浪费 (即“节流”)。同时, 大量引入可再生绿色能源 (即“开源”), 通过能量流与信息流的智能适配, 大幅降低电网的能耗。进一步地, 通过网络功能虚拟化、通信与计算资源的高能效协同, 以及移动智能体的分布式计算与协同等手段, 实现绿色计算与人工智能算法。

**关键词:** 绿色通信; 可再生能源; 超蜂窝网络; 人工智能; 绿色计算

**中图分类号:** TN929.5

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2022041

## Green communication and networking for Carbon-peaking and Carbon-neutrality: challenges and solutions

NIU Zhisheng<sup>1,2</sup>, ZHOU Sheng<sup>1,2</sup>, SUN Yuxuan<sup>1,2</sup>

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

2. Beijing National Research Center for Information Science and Technology, Beijing 100084, China

**Abstract:** Driven by the Carbon-peaking and Carbon-neutrality strategic goals, future mobile communication and networks need to drastically reduce energy consumption while satisfying the growing traffic demand. Therefore, it is necessary to study the theories and technologies used to send more information bits with less energy (SMILE). To meet the challenge, it was far from enough to rely solely on the improvements of wireless transmission technologies and hardware. It was necessary to explore efficient energy utilization mechanisms from the perspective of the system and network. Both energy saving and renewable energy utilization approaches were proposed, and solutions for the growing computing energy consumption were provided. Specifically, by introducing hyper-cellular network, agile coverage and elastic access were realized, and thus service base stations and edge servers could enter the sleeping mode to save energy when the traffic volume was low. Meanwhile, by the introduction of renewable energy and the smart adaption of energy flow and information flow, the energy consumption of the power grid could be greatly reduced. Furthermore, green computing and artificial intelligence could be realized through network function virtualization, energy-efficient coordination of communication and compute resources, and distributed computing and coordination approaches of moving agents.

**Keywords:** green communication, renewable energy, hyper-cellular network, artificial intelligence, green computing

### 0 引言

各种信息技术, 特别是遍布于地球各个角落的

通信与网络基础设施, 在经历了几十年高速发展之后都逐渐遇到了能耗的瓶颈, 其快速增长的自身能耗也已演变为全球气候变暖的因素之一。以移动

收稿日期: 2022-01-12; 修回日期: 2022-02-09

基金项目: 国家重点研发计划基金资助项目 (No.2020YFB1806605)

Foundation Item: The National Key Research and Development Program of China (No.2020YFB1806605)

通信网络为例,随着业务量的迅猛增长,移动通信已成为高能耗产业,其中 80% 的能耗在基站侧。例如,据《中国移动通信有限公司可持续发展报告 2018》,该公司同时运营的 350 万台 2G、3G、4G 基站消耗电力 247 亿千瓦时,占总运营成本的 60% 以上。2019—2021 年,随着 5G 网络的建设,该公司又增加了近 70 万台 5G 基站,且每台 5G 基站的功耗为 4G 基站功耗的 3~4 倍,因此网络总能耗还在持续上涨。进一步地,未来 6G 网络需要大量引入计算能力与人工智能算法,为此需要额外大量部署边缘服务器,而一台典型边缘服务器的功耗可达 2 kW 以上,其能耗水平也是非常可观的。可见,随着移动通信向 5G 和 6G 时代的迈进,网络能耗将显著增加,严重制约移动通信的可持续发展。

众所周知,能源危机与环境污染问题已经成为制约人类社会可持续发展的 2 个主要瓶颈。中国已经制定了“建设资源节约型和环境友好型社会”以及“2030 年碳达峰、2060 年碳中和”的“双碳”战略。因此,迫切需要研究并突破可在提高移动通信系统容量的同时大幅降低全网能耗的理论与技术,构建能效与资源联合优化的大容量、高可靠和低时延的移动通信系统。

显然,以提高能量效率(单位能量所能传送的信息量)为主线的传统意义上的绿色通信并不足够,因为各种预测均显示,随着全息媒体和万物互联时代的到来,未来 5~10 年人们对网络容量的需求仍将呈现指数增长的态势,因此单纯提高能量效率并不一定能够降低总的能量消耗。由此引出了一个非常基础性的科学问题:能否以及如何使用更少的能量来传送更多的信息,即 SMILE (send more information bits with less energy)?

为了应对该挑战,仅靠无线传输技术的改进和硬件实现水平的提高是远远不够的,因为经典信息理论,如指导移动通信 70 余年发展的香农定理告诉我们:要想提高传输容量,要么需要更大的带宽或是更大的时间与空间自由度,要么需要大幅度提高信噪比,即提高发射功率。尽管毫米波、太赫兹等高频段可以提供更大的带宽,但其巨大的传输损耗以及由此引发的频繁越区切换,再加上超高频信号处理的代价,将在很大程度上抵消大带宽所带来的好处,频谱资源只能是越来越受限,时间与空间自由度的获取也越来越受限于站址资源及信号处理,因此提高发射功率成为最终手段。

本文跳出物理传输层的范畴,从系统和网络的角度探索能量的高效利用机理与方法。具体地,本文从能量的“节流”和“开源”2 个维度展开,引入能量流与信息流智能匹配的机理,即 TANGO (traffic-aware network planning and green operation)<sup>[1]</sup>,并针对日益增长的计算与人工智能算法能耗给出解决方案。首先,通过引入超蜂窝网络架构实现网络的柔性覆盖与弹性接入,使业务基站和边缘服务器在业务量较低时可以进入休眠状态,减少能量的浪费(即“节流”)。同时,大量引入可再生绿色能源(即“开源”),通过能量流与信息流的智能适配,大幅降低电网的能耗。进一步地,通过网络功能虚拟化、通信与计算资源的高能效协同以及移动智能体的分布式计算与协同等手段,实现绿色计算与人工智能算法。

本文分别论述了基于柔性覆盖与弹性接入的超蜂窝网络、可再生能源供电的绿能通信与网络以及绿色计算与人工智能算法所面临的技术挑战,并给出典型的解决方案。

## 1 基于柔性覆盖与弹性接入的超蜂窝网络

### 1.1 控制覆盖与业务覆盖适度分离的超蜂窝网络新架构

蜂窝架构是移动通信系统提高频谱效率最有效的手段,通过不断地缩小蜂窝小区或小区分裂即可大幅度提高网络容量,成功地支撑了过去 40 多年来移动通信的飞速发展。但进一步缩小蜂窝小区不仅会增加网络覆盖的成本,而且会引入严重的小区间干扰,导致网络运行成本和能耗成本大幅攀升。与此同时,未来 5G 及其后续演进还需要针对物联网应用提供超大链接、超低时延和超高可靠的服务,现有硬性覆盖的蜂窝架构难以应对。因此,如何将现有蜂窝架构改造得更加绿色和更加智能是移动通信可持续发展亟待解决的关键科学问题。

现有蜂窝体系架构及其演进系统都是以基站为中心的,兼备网络覆盖、信道测量、用户接入控制以及实际数据传输等多重身份,因此即使网络负载很低,甚至完全没有用户业务传输需求时,所有基站也都需要全时开启,以保证网络的覆盖,这实际上是现有蜂窝网络能耗迅速增加的主要原因。

为此,针对未来移动通信发展所面临的频谱与能耗瓶颈,本文提出了一种基于控制覆盖与业务覆盖适度分离的超蜂窝(hyper-cellular)网络架构<sup>[2]</sup>,如图 1 所示。它超越了传统蜂窝架构的局限,通过永

远在线的控制覆盖与按需部署的业务覆盖的适度分离，引入网络的柔性 (agile) 覆盖、资源的弹性 (elastic) 匹配以及业务的匹配 (adaptive) 服务新机制，即网络的信号覆盖、资源匹配以及服务模式均随业务需求以及网络状态的变化而动态适配，使负载较低的业务基站可以按需进入休眠状态，从而大幅提高基站的资源利用率，降低全网的能耗，实现能效与资源的联合优化。

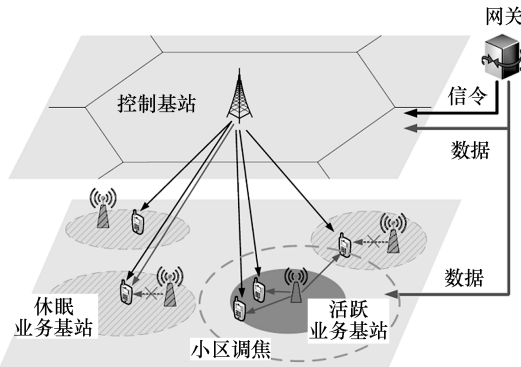


图 1 基于控制覆盖与业务覆盖适度分离的超蜂窝网络架构

具体地，超蜂窝网络架构的核心是控制覆盖与业务覆盖可按需在时域和空域上独立进行动态调整的智能网络。控制覆盖用于管理用户的接入请求以及低速率网络控制信号的传递，可使用大区覆盖模式，相对固定，是网络的骨架；业务覆盖用于为用户提供高速数据的传输，按需部署，灵活而高效。相应地，基站也分为两类，一类为控制基站 (CBS, control base station)，提供大范围的控制覆盖，实现和用户终端间的控制信令传输；另一类为业务基站 (TBS, traffic base station)，提供小范围的业务覆盖，为用户终端提供高速数据传输业务。如图 1 所示，同一个控制基站覆盖范围内有多个业务基站，可以根据业务量的大小进行业务基站间的动态资源调度和基站休眠，从而降低系统总能耗并提高资源利用率。可见，超蜂窝网络可以在保证网络无缝覆盖和频谱效率的同时大幅度降低整体能耗，从而解决了传统蜂窝架构的绿色可持续发展问题。

该解决方案的核心是能量的“节流”，即通过柔性覆盖减少蜂窝网络中的能量浪费，解决了传统蜂窝架构的绿色可持续发展问题。它不同于所谓“无蜂窝”的概念，仍然沿用了蜂窝覆盖的基本架构，但实现了按需部署与使用，体现了根据业务实际需求部署与运营网络的柔性思维，即 TANGO 的理念。这种全新的超蜂窝网络架构是自 70 多年前

蜂窝网络概念被提出以来的大变革，为未来移动通信 (包括 5G 及其后续演进) 网络的长远发展开辟了一条可持续发展的道路，同时保持了蜂窝网络的基本理念，保障了网络发展的后向兼容性。实际上，目前 3GPP R16 中的 CU/DU 分离的概念基本上等效于超蜂窝架构，5G 标准中的 NSA (non-standard-alone) 模式也可以视为超蜂窝架构的变形。

## 1.2 超蜂窝网络的网络能效理论

进一步地，在超蜂窝架构下，需要研究网络能效与网络覆盖能力、频谱效率、用户分布、业务特性等的理论关系，即在给定频谱资源和业务覆盖需求下的最小能量理论 (简称为“网络能效理论”)，包括：如何以最小的能量实现所需的网络覆盖及业务服务需求，网络能效与频谱效率之间存在什么样的理论关系，业务的随机性和动态特性究竟会如何影响网络能效？如果只考虑单链路的发射功率，经典信息论则告诉我们：链路能效与频谱效率之间存在单调的折中关系。但如果额外考虑信息处理等能耗，则这种折中关系就会变得非常复杂，至少不再单调。进一步地，如果考虑全网的能量效率与端对端时延分布，则两者的关系会变得更加复杂，甚至可能不再是折中关系。

首先，考虑网络能效与业务时延之间的理论关系。为了提高网络能效，无论是引入柔性覆盖、弹性接入，还是引入匹配服务的机制都可能会给用户带来额外的时延。如何在用户可以容忍的时延下获得尽可能多的能量节省，牺牲一定的时延性能可获得多少能量节省等，还是亟待解决的基础理论问题。具体地，本文针对超蜂窝网络建立了一种基于休眠的排队模型，分析了迟滞休眠机制、单重休眠唤醒机制、多重休眠唤醒机制和累积任务唤醒机制的性能，给出了网络能效与业务时延及其抖动的闭式关系，以及启动时间、休眠时间、累积任务数和迟滞时间等对系统性能的影响，得到了多组参数影响下能效-时延关系的闭式解<sup>[3]</sup>。

超蜂窝网络的能效时延关系如图 2 所示，其中， $N$  为唤醒基站的负载阈值； $P_t$  为发射功率； $C^2$  为间歇泊松过程 (IPP, interrupted Poisson process) 或交互泊松过程 (SPP, switched Poisson process) 业务的方差系数，其值越大表征业务的随机性越大。数值分析结果表明 (图 2 中虚线所示)，在考虑电路功耗与基站休眠的情况下，能耗与时延并不总是呈现折中关系，两者之间可能存在“共赢”区间，即降低全网能耗不一

定要牺牲业务时延，通过合理设计系统及休眠参数，完全可以同时实现能耗与时延的降低。当进一步掌握业务的统计分布时，则可以将功率控制与动态休眠进行联合优化，从而获得最优的能耗与时延折中（图 2 中实线所示）。分析还发现，基于累积负载阈值（图 2 中的参数  $N$ ）的休眠策略比基于休眠时长的策略具有更好的能耗-时延折中。这些结论明确了实时监测网络负载状态对休眠决策的重要性，为控制与业务分离的蜂窝网架构演进提供了理论依据。

其次，考虑网络能效与业务随机性之间的理论关系，引入 IPP 和 SPP 以描述业务到达过程的抖动以及自相关性。研究给出了 IPP 模式下网络能效与业务时延的闭式解<sup>[4]</sup>，业务突发性和业务非均匀性对能效的影响分别如图 3 和图 4 所示，图 4 中  $\alpha_1$ 、 $\alpha_2$ 、 $\alpha_3$  分别为热点小区、热点邻小区、热点外围小区的归一化业务强度。分析结果表明，业务突发性（抖动）越大，网络能效反而越高。这打破了传统的认知，即在服务者（基站）没有休眠机制的排队系统中，业务突发性永远会带来排队性能的损失。同样地，针对业务在空间域上的非均匀分布特性，通过马尔可夫决策理论分析，发现了“业务分布越不均匀，基站动态休眠所带来的节能增益越大”的结论<sup>[5]</sup>。该结论也是超越传统认知的，表明负载均衡并非永远有效，在基站可以休眠的网络中，应适时进行负载集中<sup>[6]</sup>，增加低负载基站的休眠机会，从而提高全网的能量效率。

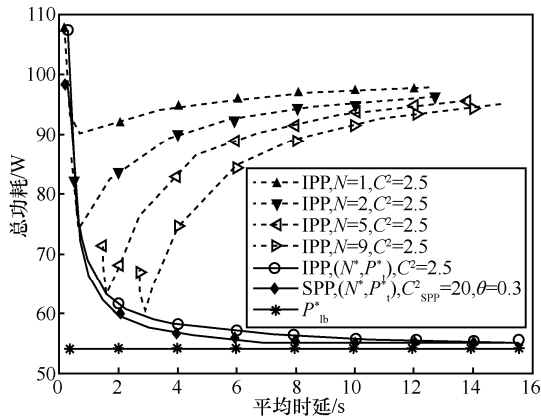


图 2 超蜂窝网络的能效时延关系

具体地，为了充分挖掘业务突发性及非均匀性所带来的休眠机会，文献[6]提出了定向负载转移机理，改变了负载均衡的传统理念。特别地，通过负载聚合，主动改变业务的空间分布，将不同空间位置和异构网络中不同层之间的业务通过卸载进行整合，由少数基站集中进行服务，以提供更多的基站休眠机会。

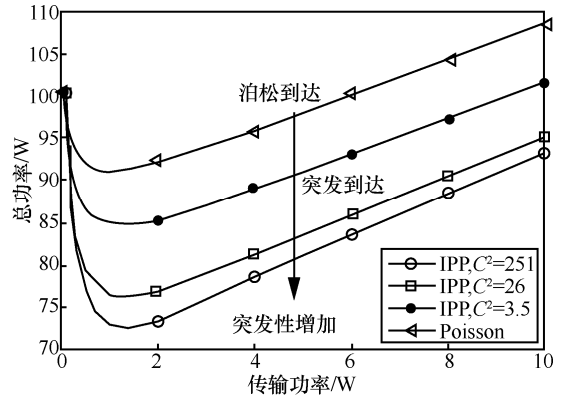


图 3 业务突发性对能效的影响

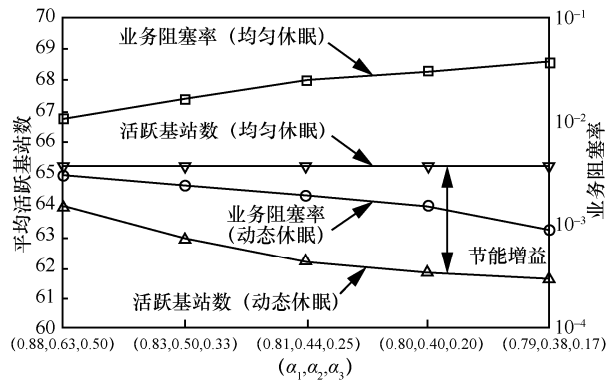


图 4 业务非均匀性对能效的影响

如果进一步考虑频谱资源的最优分配，则在基站可进行动态休眠的超蜂窝网络架构下，基站的工作状态（休眠与否）、业务负载和无线资源分配之间紧密耦合，传统无线资源分配机制将难以适用。文献[7]运用随机几何理论，给出了最优的宏/微基站频谱重用因子、基站休眠概率和业务负载三者间的闭式理论关系，即

$$\lim_{W \rightarrow \infty} \beta^* = \frac{\Lambda + c\rho_M + \rho_m}{c\Lambda + \rho_m + c\rho_M} \quad (1)$$

其中， $\beta^*$  为最优的宏/微基站频谱重用因子，即微基站所分配带宽与宏基站所分配带宽的比例； $\Lambda$  为用户密度（表征业务负载）； $\rho_M$  和  $\rho_m$  分别为宏基站和微基站密度（表征休眠概率）； $c$  为由宏微基站发射功率和路径衰落因子决定的系统参量； $W$  为系统可用带宽。基于此关系设计的基站动态休眠和频率重用算法典型场景下可节省 50% 的网络能耗。

### 1.3 超蜂窝网络业务基站的最优休眠控制

本节进一步考虑超蜂窝网络业务基站的最优休眠控制。针对泊松到达业务和单一服务者，已有研究证明了最优休眠控制具有双门限结构<sup>[8]</sup>，即队列长度低于休眠门限值时基站进入休眠状态，但只

有在队列长度大于开启门限值（该值一般大于休眠门限值）时基站才开启服务。但针对突发业务到达以及多服务者系统而言，基站的最优休眠策略还有待研究。文献[9]首先考虑间歇泊松过程以及交互泊松过程 2 种突发业务到达单一基站情形，并考虑基站休眠与开启的能耗代价，基于马尔可夫决策理论严格证明了最优休眠控制策略仍然具有双门限结构，但会额外增加一个观望（wait-and-see）特性，即队列长度在低于休眠门限值时并非立即进入休眠状态，而是观望一段时间后才进入休眠状态；反之，针对开启门限也是同样，且观望时长随业务到达突发性的增大而加大，即基站的休眠与唤醒应更加保守。最优休眠控制的双门限结构和观望特性分别如图 5 和图 6 所示。进一步考虑突发业务到达多个基站的情形，文献[10]理论证明了最优休眠策略具有多门限结构，且具有观望特性。

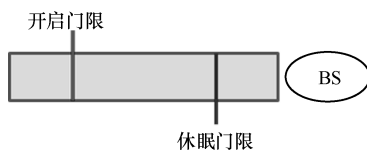


图 5 最优休眠控制的双门限结构

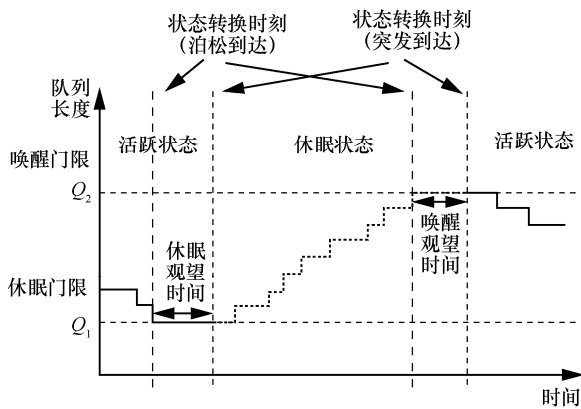


图 6 最优休眠控制的观望特性

文献[3]研究了在平均时延约束下的功率最优基站休眠策略。在单重休假、多重休假及累积负载阈值机制下，证明了时延约束均存在一个阈值  $\tau_{th}$ ，使最优休眠控制具有以下结构。1) 在  $\tau < \tau_{th}$  时，最优休假时间  $h_V$ （单重、多重休假机制）或累积负载阈值  $N$ （累积负载阈值机制）对不同的  $\tau$  保持不变，而最优的休眠概率  $p_v$  随  $\tau$  单调增加，即此时只需调整休眠前的迟滞时间即可。2) 在  $\tau \geq \tau_{th}$  时，最优休眠概率  $p_v$  恒为 1（即最优迟滞时间恒为 0），而最优休假时长  $h_V$  或最优累积负载阈值  $N$  则随  $\tau$  单调

增加，举例来说，累计负载阈值策略下的阈值  $\tau_{th}$  满足<sup>[3]</sup>

$$\tau_{th} = \bar{T}_0 + \frac{\lambda h_s^2(1+c_s^2)+2h_s}{2\lambda h_s+2} \quad (2)$$

其中， $\bar{T}_0$  是没有基站休眠条件下的时延（基于 M/G/1 队列建模）， $\lambda$  是业务到达率， $h_s$  和  $c_s^2$  分别是启动时间的均值和方差系数。

以上分析对业务请求的到达模型均做了某种假设，但在实际系统中业务的到达过程可能非常复杂，难以用经典概率模型表述。为此，本文将强化学习与业务模型的先验知识相结合，提出了对业务模型与最优休眠策略同时学习的在线算法<sup>[11]</sup>。通过现网业务数据的测试，表明所提在线学习算法相比基于模型的策略和单纯强化学习算法，在时延和能耗的折中性能上均有明显优势。

## 2 可再生能源供电的绿能通信与网络

为了实现绿色移动通信，除了以网络协作等方式进行系统中能效和资源优化之外，还可以通过引入能量收割技术利用可再生能源（如太阳能、风能等）对通信系统供电，进一步降低传统能源的使用，从而降低碳排放。这相当于能量“开源”的思路，某种意义上讲是一种“取之不尽、用之不竭”的新能源。同时，由于不需要接入电网，可再生能源供电的通信设备可以进行灵活的部署，并且相比于电池供电的设备有更长的寿命，具有自我持续性。

由于上述优势，可再生能源在通信系统中的应用受到了越来越多的关注。2009 年，国际标准化组织第三代合作伙伴计划（3GPP, 3rd Generation Partnership Project）对各种可再生能源在通信网络中的使用场景、使用方法和可行性提出了标准化建议 TR 102.532，中国通信标准化协会对使用可再生能源的基站供电也进行了标准化工作，中国移动在西藏部署的基站中约  $\frac{2}{3}$  使用了可再生能源供电。随着智能电网技术和电池储能技术的飞速发展，可再生能源所产出能量的存储和利用效率也在大大提高，能量收集技术已步入实用阶段。据国际能源署统计，2020 年可再生能源电力占比快速提升至近 30%。我国可再生能源发展水平与国际水平并驾齐驱，2020 年可再生能源发电量达到 2.2 万亿千瓦时，占全社会用电量的 29.5%。可以预见，未来信息系统的供电模式将日趋多样化，甚至会出现完全依赖

于可再生能源的自供电信息系统,这不仅可以大幅降低信息系统对传统电网的依赖,而且可以减少布线需求和运维成本,同时也契合“碳中和”的战略方向。特别是针对基站密集部署的超密集组网(UDN, ultra-dense network)和分布式多输入多输出移动通信系统,其广泛部署的小基站或天线前端很有可能无法或难以直接连接到电网,需要完全依赖于可再生能源供电。

然而,可再生能源的引入给通信系统带来了许多新的问题和挑战,这些挑战主要来自可再生能源的不确定性。在传统电网供电的移动通信中,能量供给是稳定的,研究人员主要面对的是业务需求的随机性和传输信道的随机性问题。而在可再生能源供电的移动通信系统中,除了上述 2 种随机性之外,需要同时面对能量供给的随机性。传统资源分配手段面向吞吐量最大化而贪婪式的使用能量,在可再生能源供电条件下,会严重影响用户服务质量。

为此,需要扩展传统绿色通信的内涵,即仅仅关注能量效率是远远不够的,需要同时关注能量持续性指标。该指标包含能量中断和能量溢出 2 个层面。前者指虽有信息服务的需求但却没有足够能量供给所导致的服务中断现象;后者指由于能量缓存容量的限制,虽有充足的能量供给但却没有相应的信息服务需求而导致的能量溢出现象。很显然,这 2 种现象都是需要尽量避免的,尤其是服务中断现象。

上述问题基本上可以转换为一个信息流如何与能量流更好匹配的问题,即 TANGO<sup>[1]</sup>思维。一方面,可以让能量流更好地适配信息流,即在需要传输更多信息时分配更多的能量;反之,则分配更少的能量。这在传统电网供电的信息系统中是比较容易实现的,但在可再生能源供电的信息系统中将面临巨大的挑战。虽然可以通过调控能量缓存器(即电池)的充放电来解决,但如何做到时间尺度的实时匹配以及空间维度的动态调度是问题的难点。另一方面,可以让信息流更好地适配能量流,即能量供给比较充足时可以传输更多的信息,比如引入缓存与推送机制,通过对用户需求的预测计算出用户的潜在需求,在能量充足时进行推送或缓存;反之,在能量供给不足时则可以减少信息的传递,比如只提供最基本的信息服务,而推迟提供增强服务,直至能量充足,或是将任务卸载到能量充

足的基站。

为此,针对未来移动通信网络的能量供给高度动态可变场景,亟待研究以下问题:1) 移动通信覆盖保障的机理与柔性覆盖理论;2) 无线频谱与能量资源高效利用理论;3) 业务服务的智能推送与缓存机理与方法等。从而揭示能量状态信息(即能量到达的统计特性和电池状态)对优化无线资源管理、传输调度与用户服务的价值,刻画动态能量供给下的无线链路和网络性能,以及能量存储与信息存储之间的互换机理等。

下面给出信息流与能量流适配的 3 个例子。

## 2.1 基于电池储能信息的能量与信道适配

为了定量描述如何根据电池储能状态捕捉传输机会,文献[12]针对收发端均由可再生能源供电的无线链路,利用电池储能信息优化发端的发射功率和收端的接收行为,最小化由于信道衰落或能量不足导致的传输中断概率。具体地,在电池容量较大的条件下,在所有在线功率控制策略(即没有未来能量到达信息,只有历史信息)中,提出并证明了最优功率控制策略是基于电池状态的门限形式,即

$$P_S^t = \begin{cases} P_S, B_S^t \geq P_S \Delta \\ 0, \text{其他} \end{cases} \quad (3)$$

当  $t$  时刻发端电池储能  $B_S^t$  大于门限值  $P_S \Delta$  时,应以固定发射功率  $P_{tx}$  传输,满足发端整体功率为  $P_S = (1 + \alpha)P_{tx} + P_{C,S}$ , 其中,  $\alpha$  是与功放效率有关的常数,  $P_{C,S}$  是发端的电路功耗,  $\Delta$  是时隙长度;否则应停止发送,等待能量累积。当收发端不共享电池状态时,最优门限  $P_S^*$  的闭式解为

$$P_S^* = \max\{\lambda_S, B_{th}\} \quad (4)$$

其中,  $\lambda_S$  是发端能量到达率,参数  $B_{th}$  为

$$B_{th} = \frac{1}{2} \left[ 2P_{C,S} + \delta + \sqrt{\delta(4P_{C,S} + \delta)} \right] \quad (5)$$

其中,  $\delta$  是一个与传输速率需求和信道增益相关的常数。可见,最优门限  $P_S^*$  与平均信道增益、能量到达率和电路功耗之间存在较强的关联关系。

当收发端可以共享电池状态时,文献[12]已证明最优策略依然是基于电池状态的门限形式,即仅在收发端均有足够能量进行接收和发送的时候才传输,即

$$P_S' = \begin{cases} P_S, & B_S' \geq P_S \Delta \text{ and } B_D' \geq P_D \Delta \\ 0, & \text{其他} \end{cases} \quad (6)$$

其中， $B_D'$  是收端电池储能状态， $P_D$  是收端整体功率。此时最优门限  $P_S^*$  的闭式解为

$$P_S^* = \max \left\{ \lambda_S, B_{in}, \frac{\lambda_S P_D}{\lambda_D} \right\} \quad (7)$$

其中， $\lambda_D$  是收端能量到达率。可见，在收发端共享电池状态时，最优门限  $P_S^*$  与接收端的能量到达率和接收电路功率有关。收发端均由可再生能源供电的无线链路中断概率如图 7 所示，其中，横坐标为 100 mW 归一化的能量到达率。从图 7 可以看出，共享电池状态 (Joint) 可消除收发端能量获取的随机差异，相比于不共享电池状态 (Disjoint 和 Linear) 能显著降低传输中断概率，尤其当能量到达率较低时。

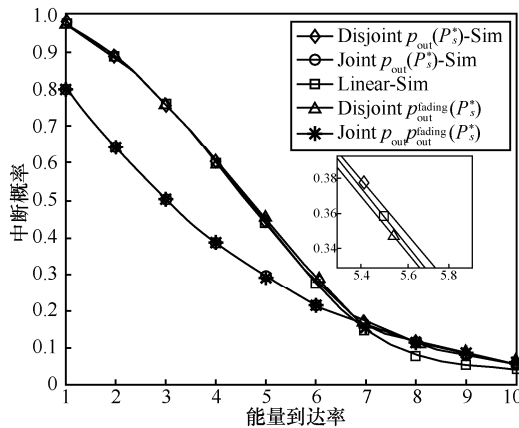


图 7 收发端均由可再生能源供电的无线链路中断概率

为了通过电池协作管理使用可再生能源供电和电网，保障链路性能，文献[12]针对发端可再生能源与电网混合供电的衰落信道，发现并证明了吞吐量最大化问题的最优功率分配具备两阶段注水结构，即第一阶段基于电池容量约束定向注水所采集的可再生能源电量，第二阶段在现有水面上用传统注水法分配电网功率。上述结论简洁且物理意义明确，是提出可再生能源与电网协同的最早论文之一，拓展了前人学者针对仅有可再生能源供电条件下的最优功率控制结论。文献[13]则进一步将上述结果拓展到了多天线场景，发现并证明了最优的空时联合注水结构。

## 2.2 基于负载转移的能量与业务空间适配

移动业务不仅在时域上动态变化，在空间域的分布也是高度不均匀的。传统上为了保障任何时间、任何地点的业务覆盖，往往采用最坏情况设计

的原则，即在任何地点均按照可能出现的峰值流量部署网络资源。这会不可避免地造成网络资源的浪费，因为在一天之中某个基站的流量能够真正达到峰值的时间是非常短的。

在超蜂窝架构下，由于基站可以随时随地进入休眠状态，因此即使按照峰值流量来配置，网络资源浪费也不会很大，同时也不再有必要保持各基站之间的负载均衡，完全可以通过适当的负载转移将低负载基站的业务转移至相邻的中高负载基站中，通过关断低负载基站节省能量的同时，提高中高负载基站的资源利用率。当然，针对传统蜂窝网络，在负载较大时，通过负载转移实现基站间的负载均衡还是会提高整体服务质量的。

为此，本文提出了一种将业务负载转移与主被动基站休眠，以及多基站协作传输联合优化的策略<sup>[14]</sup>，等效地实现了能量在基站间的“转移”，解决了能量到达及其存储在基站间的不均匀性与业务在小区间动态变化相失配的问题。具体地，通过将负载转移到能量到达或存储较充分的基站进行服务，实现了能量在基站间的等效“转移”。特别地，区别于贪婪式的用能方式，采用动态规划理论，体现当前时刻的基站休眠和资源分配决策对未来电池中可用能量的影响，减少了能量不足可能导致的服务中断。

进一步地，文献[6]将上述工作拓展至异构蜂窝网，即同时考虑基站类型的异构与能量供给方式的异构，通过不同层之间的业务卸载，在空间上实现了能量与业务的适配。同时，利用多基站协作，联合优化多基站联合预编码和可再生能源供电，从协作传输的角度实现了基站间的能量“转移”。面向能效优化的异构蜂窝网络负载聚集如图 8 所示，仿真表明，低负载条件下进行负载聚集（而非负载均衡），可显著降低网络能耗。

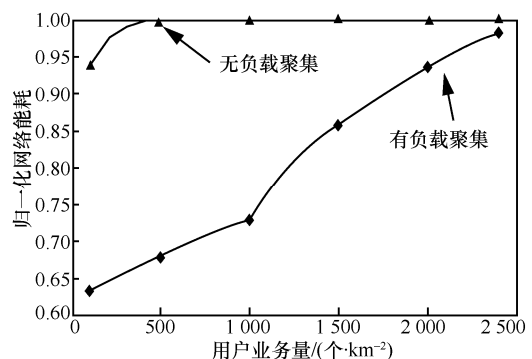


图 8 面向能效优化的异构蜂窝网络负载聚集

### 2.3 基于缓存推送的能量与业务时域适配

针对可再生能源供能的随机性与业务请求在时间上不匹配的问题,提出了一种基于内容缓存与提前推送的可再生能源有效利用机制<sup>[15]</sup>,其核心思路是:当电池电量充足或可再生能源供能充分时,在基站处主动缓存受关注度高的内容,或将已缓存内容提前推送至用户端,使用户仍可在基站电池电量或可再生能源供能不足时从本地获取所需的信息。这一思路克服了可再生能源供能的随机性以及电池容量约束,使能量可以“提前”使用以适配业务在时间上的变化,等效于将能量在时间上进行了“转移”。文献[16]中基于马尔可夫决策理论给出了最优的推送策略,提出了基于电池状态的门限推送策略,该门限值与未推送内容的访问概率相关,并证明了该门限策略在电池容量较大时或能量到达率较低时可逼近最优策略。基于电池状态的推送缓存策略的性能如图 9 所示,其中  $\lambda_c$  为内容的更新率。数值仿真表明,所提策略的性能关于电池容量稳健,即较小的电池容量也不会对性能产生明显的影响。

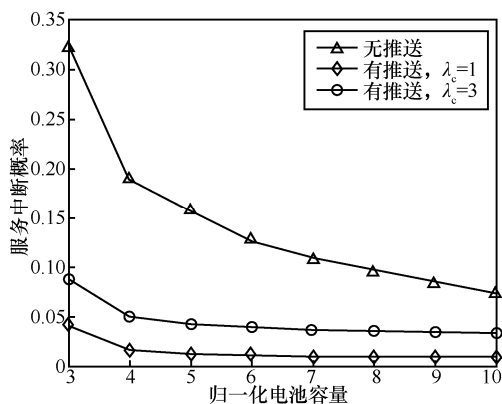


图 9 基于电池状态的推送缓存策略的性能

## 3 绿色计算与人工智能算法

另外一种可能的“开源”途径是引入更多的计算与存储资源。随着算力的提升以及软件定义网络(SDN, software-defined network)、网络功能虚拟化(NFV, network function virtualization)等技术的发展,无线通信系统中的信号和基带处理等已开始逐步采用通用计算和存储资源<sup>[17]</sup>,以支持灵活的网络重构与切片,并通过基带信号处理的集中简化基站前端的功能,从而降低全网成本与能耗,是下一代蜂窝系统的重要使能技术之一。某种意义上讲,相当于用计算资源来换取通信资源的节省。然而,经

典云化无线接入网(CRAN, cloud radio access network)过度依赖于前传通信资源进行大规模的高带宽基带信号汇聚,造成通信与计算资源的失配,从而降低资源的利用效率。因此,如何实现通信与计算资源的协同设计与优化,提高网络整体的效能,是一个亟待解决的科学问题。

与此同时,随着移动互联网、大数据和人工智能等技术的快速发展,无线通信网络将不仅要承担大容量低时延的通信服务,还需要满足各类无线节点随时随地计算并共享计算结果的迫切需求。具体地,从物联网传感器到智能手机,从基站设备到网联汽车,未来都将具备可定制或通用的计算资源,从而改变原有的单一计算模式,并通过无线网络实现协同计算,支持自动驾驶、智能制造、智慧医疗、安防和环境监测等领域的广泛应用。可见,未来无线网络将承载大量由各类机器产生的、以计算和控制业务为主体的全新业务,使通信与计算在需求和资源上均存在强耦合。换言之,无线网络将从原来搬移数据的单纯管道进化为面向机器学习和协同控制的信息感知、传输、处理为一体的智慧网通道。因此,如何协调异构移动智能体,实现分布式学习与决策是另一个亟待解决的科学问题。

当然,引入算力与人工智能算法都需要付出巨大的能耗代价,因此算力与人工智能算法的部署与运营也应该遵循柔性的思维,只在必要时才启用,即 AI-on-Demand 或简称为 Green AI。具体地,算力与人工智能算法的使用应至少满足以下 4 个条件。

1) 建模不再可能或是模型非常复杂。例如,未来毫米波甚至太赫兹波段的信道模型以及阻挡模型,截至目前还未找到比较好的数学模型,而且有迹象表明,即使能够建立数学模型,其复杂度也将是非常高的,难以进行数学分析。

2) 高质量的数据可以获取且数据量足够。众所周知,AI 算法的可靠性严重依赖于高质量且足够数量的训练数据,传统语音、图片、视觉等数据比较容易获取,且通用性很强。但通信信道、通信业务、网络环境等通信相关的数据一般都与数据获取的时间和地点紧密相关,且随时间和地点不断变化,因此一般通用性很差,必须随时随地反复获取和更新。

3) AI 算法不是很复杂,使基于 AI 的决策可以实时进行。利用 AI 算法的最终目的是做决策,而

决策往往要求非常高的实时性，否则难以应用于诸如虚拟与增强现实、自动驾驶等场景。

4) 基于 AI 决策的收益或性能增益大于所需付出代价（包括能耗代价）。众所周知，基于 AI 的决策往往需要训练规模很大的深度神经网络，并维护大规模的数据库，其代价（特别是能耗代价）非常高。因此基于 AI 的决策所能带来的收益必须足够大，否则没有意义。

针对上述科学问题，下面给出 3 个典型的解决方案例。

### 3.1 基于网络功能虚拟化的软件定义超蜂窝网络

如 2.1 节所述，超蜂窝网络可以在保障蜂窝网络无缝覆盖和频谱效率的同时，通过引入基站的动态休眠和资源调度，大幅度降低网络整体能耗。但业务基站的密集部署仍将面临巨大的部署与管理成本，而且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。

为此，结合当前移动接入网的发展趋势，本文将空口的控制/业务分离、软件定义网络中的控制/

用户面分离有机结合，提出了基于云架构、基站功能虚拟化的软件定义超蜂窝网络系统框架<sup>[18]</sup>，如图 10 所示，并实现了如图 11 所示的演示验证平台。该平台实现了 GSM/GPRS 和 LTE 协议栈，支持商用终端接入。实测表明，所提出的软件定义超蜂窝架构可通过灵活的控制/业务分离支持动态基站休眠，分离所带来的信令时延开销仅 0.36 ms。此外，所提的软件定义超蜂窝架构和休眠算法在不同的休眠级别下可取得约 16%~60%的能耗节省。可见，该架构可使移动网络根据实际业务需求进行动态重构变得更加容易，进一步提高了超蜂窝网络的灵活性和智能化水平。它不仅可以满足未来无线移动通信在谱效、能效、用户体验和弹性接入等多重指标的综合需求，还可支持灵活的无线网络重构、剪裁及网络升级等，与核心网的 SDN 取得良好的协调。

进一步地，基于虚拟化基带功能的灵活分割和部署，文献[19]实现了业务需求与异质异址的无线和计算资源的动态适配，提出了高能效的计算资源虚拟化与动态调度机制，以及基于软件定义包交换的前传网全新设计<sup>[20]</sup>，其中的理论贡献如下。采用

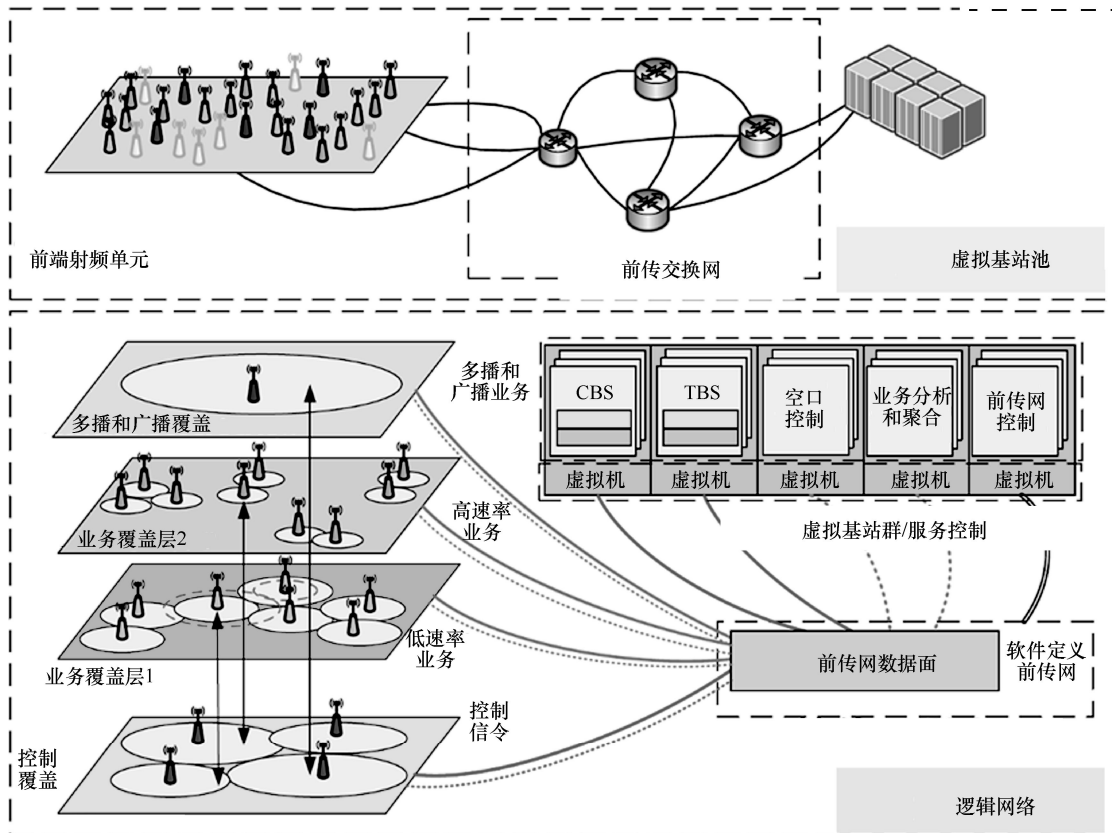


图 10 软件定义超蜂窝网络系统框架



图 11 软件定义超蜂窝网络演示验证平台

极限分析得到了虚拟基带池和包交换前传网的统计复用增益的闭式表达，并在此基础上给出了虚拟基带池规模的最优值，得出了以下两点结论：1) 适当部署计算资源可节省无线频谱和能量资源，如图 12 所示；2) 虚拟基站池规模存在一个最优值，不宜太大或太小，应采取组簇方式构建中等规模基带池，如图 13 所示，其中， $a$  为基站负载， $K$  为射频通信资源数， $P^{bh}$  为阻塞率需求，虚拟基站池规模在 50 左右即可获得主要的性能增益。

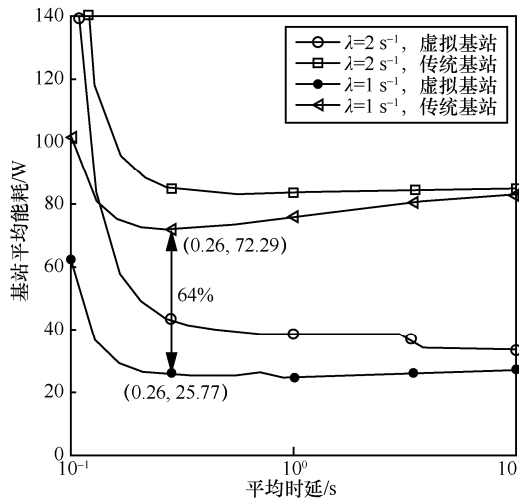


图 12 通信资源与计算资源的互换

### 3.2 通信与计算资源的高能效协同优化方法

考虑智能终端之间或智能终端与网络边缘之间通过无线传输与计算卸载，协作完成深度学习（如事件检测、视觉分析、实时决策等）的模型训练与推断任务，如图 14 所示。在此场景下，通信与计算紧密耦合。例如，在能量和计算资源受限的终端上执行复杂的神经网络推断会导致难以容忍的时延，然而将

推断完全卸载至网络边缘或其他终端将需要传输大量数据，带宽受限条件下也会增加任务完成的时延。为此，需要对无线资源和计算资源进行协同调度，通过优化计算任务在不同终端之间，以及终端和网络之间的分割与部署，提高边缘计算的能效。

如图 14(a)所示，考虑基于深度神经网络的实时推断任务由网络中的多个节点协作完成场景。将神经网络的前若干层放置在终端本地，并将该部分计算产生的中间参数由无线信道传输至其他协作终端或边缘云完成剩余网络的计算。基于神经网络剪枝技术，文献[21]提出了面向协作推断的两步压缩方法，依次对推断任务的整体计算复杂度及中间层传输数据量进行了压缩，以支持灵活的网络分割。图 15 反映了不同分割点对应的推断总时延和终端能耗，以及二者的折中关系。在 Conv<sub>1</sub> 层分割，能够以较少的终端能耗较大幅度降低推断时延，但若想继续降低推断时延，所需要的能量将大大增加。因此，在满足推断决策的时效性和精度需求的同时，应遵循“按需适度服务”的原则，优化终端和网络侧能效。

如图 14(b)所示，针对训练类任务，现有框架主要包括多终端与网络边缘联合训练深度学习模型的联邦学习 (FL, federated learning)，以及终端间协同的分布式训练 (DGD, decentralized gradient descent)。考虑联邦学习场景，文献[22]针对终端计算能力和传输性能的双重随机性，分析它们对联邦学习收敛速度的影响，基于此设计并优化终端的梯度传输调度与带宽分配，最大化模型收敛速率。针对终端能量受限场景，基于李雅普诺夫优化理论建立虚拟能量队列，以表征当前能量使用较预期的偏差。进一步设计了能量与

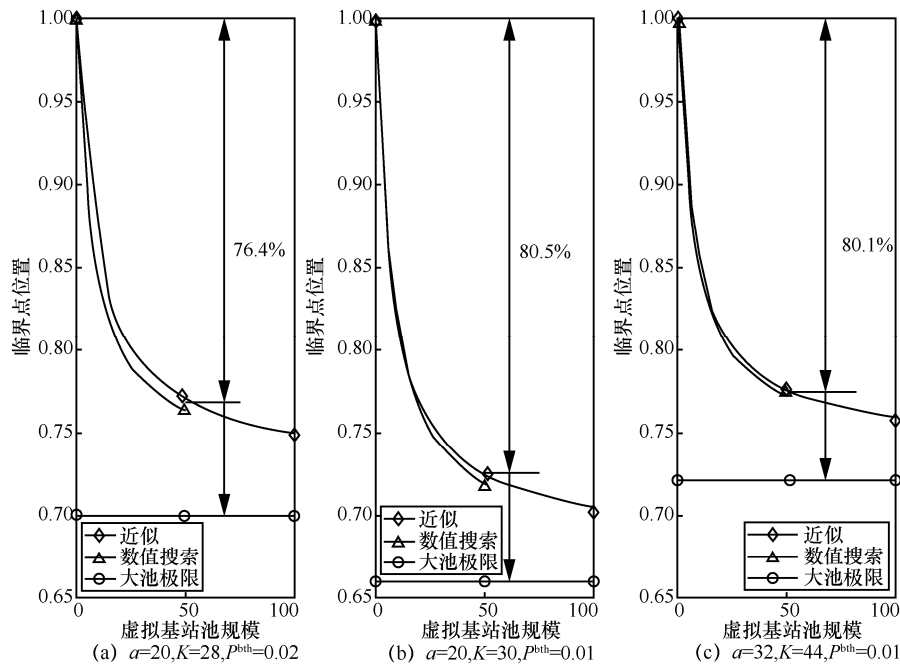


图 13 虚拟基站池规模与性能增益的关系

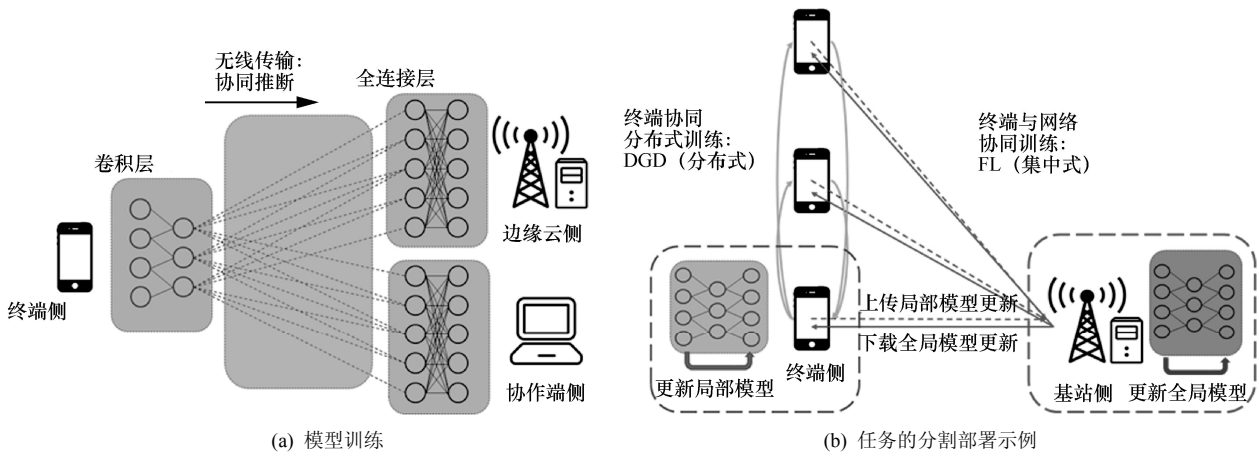


图 14 深度学习模型训练与推断任务

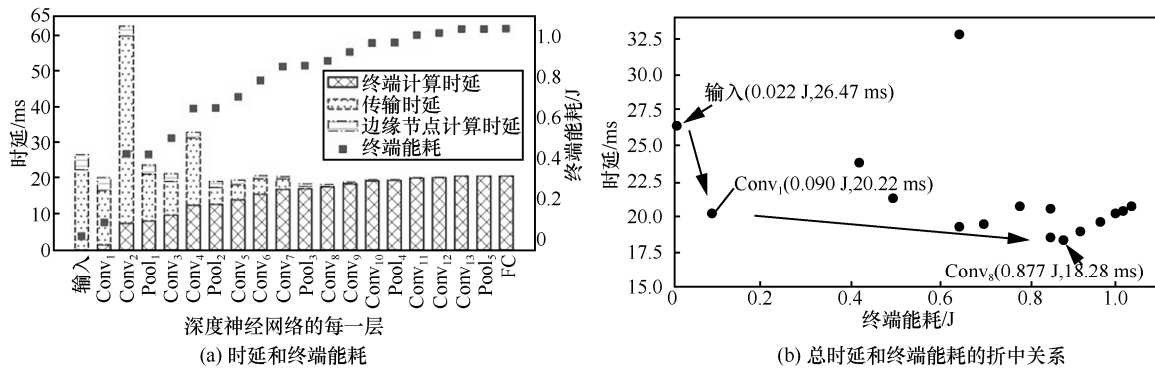


图 15 不同分割点对应的推断总时延和终端能耗及二者的折中关系

训练性能折中的效用函数, 提出终端动态调度策略, 并对该策略在能耗预测存在误差条件下的稳健性进行了分析。高效联邦学习终端动态调度策略的训练正确率和能量使用比例如图 16 所示, 本地训练数据为均质 (iid) 或非均质 (non-iid) 分布,  $m$  为每个终端的数据标签种类数目。使用 CIFAR-10 数据集的仿真结果显示, 所提动态调度策略可在满足终端能量约束的条件下, 优化联邦学习的训练性能, 在数据非均质分布的情况下, 正确率较对比策略提高 4.9%<sup>[23]</sup>。

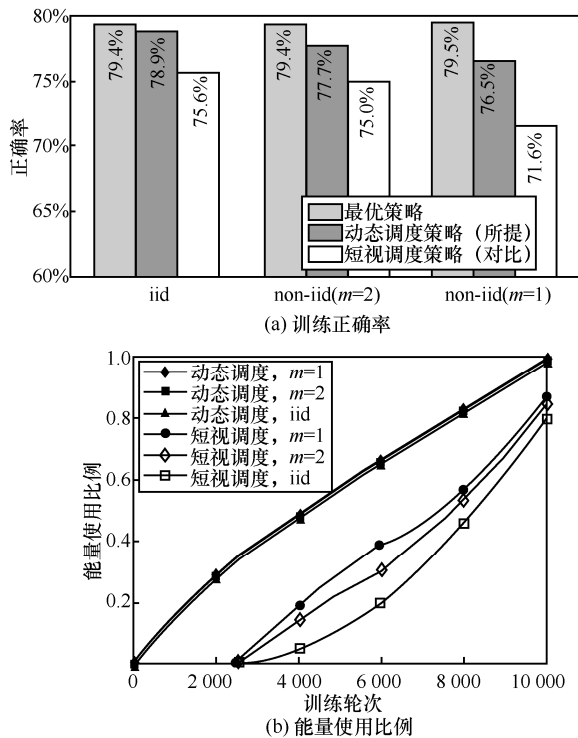


图 16 高效联邦学习终端动态调度策略的训练正确率和能量使用比例

对于多任务混合共存的场景, 可基于服务功能链刻画分割后子任务之间的依赖关系, 在终端的能耗和计算能力、频谱资源约束下, 优化任务在终端和网络之间的分割和部署, 以提供“按需适度服务”。针对边缘云的节能, 借鉴蜂窝网基站节能的思想, 提出计算任务部署的重构转移机制<sup>[24]</sup>, 在计算负载较轻时, 通过负载聚集为基站或边缘计算单元提供更多的休眠机会以实现节能。

### 3.3 移动智能体的分布式计算与协同

将终端的计算业务卸载到基站侧的通用计算资源进行处理, 可有效节约终端计算所需的能耗<sup>[25]</sup>。而用户移动会导致执行计算任务的虚拟机在基站之间迁移, 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。同时, 通信与计算异质资源的联合优化和用户调度所面临的多维系统信息获取和高复杂度优化, 在移动环境下显得尤为重要。因此, 面向网关节点的协同环境感知和自动驾驶等典型应用, 针对网络在拓扑、能量获取、频谱信道、计算资源的多维动态特性, 亟待研究如何在高动态环境下联合分配计算任务和调度无线通信资源, 在满足服务质量的同时优化全网能效。

首先, 文献[26]基于多臂赌博机理论提出了一种超密集网络下计算感知的在线切换优化机制, 综合考虑虚拟机迁移代价和信道状态, 在相同终端能耗水平下, 可以降低 20% 的计算任务完成时延。该方法不需要额外信令, 对用户移动、计算业务负载动态变化稳健, 其理论贡献如下: 传统的多臂赌博机理论只能处理固定的切换待选基站集合, 该方法可以处理时变的待选基站集合和任务大小, 且证明该方法性能与离线最优解的差距有界。

同时, 随着智能网联汽车快速发展, 未来汽车将集强大的感知、通信、计算与控制功能于一体, 并具备自供电能力。在此背景下, 本文提出了车载云计算的新思路<sup>[27]</sup>, 其架构如图 17 所示, 网络主动利用车辆的剩余算力获得机会式计算服务, 并通过车辆的移动将智能扩散到网络的各个角落。车辆的移动性可能带来更多的计算机会, 对于降低计算服务的时延和网络节能具有积极影响, 这就是所提 MEET (mobility-enhanced edge intelligence) 的概念。

针对有中心控制节点的场景, 文献[28]基于马尔可夫决策过程得到了最优任务卸载策略, 即优先卸载已复制份数最少的任务, 可最大化任务在时延约束下的成功率。车辆速度与任务完成比例的关系如图 18 所示。该研究还发现, 使任务完成成功率最大的最优车速大于传统交通流量最大化的最优车速, 这也反映了移动性对于计算服务性能的正向效应。

针对车辆间的分布式卸载场景, 文献[29]基于多臂赌博机理论, 提出了可自适应于车辆拓扑动态变化的在线学习卸载策略。为了充分利用周围车辆的计算机会, 可通过任务复制卸载的方法, 将每个任务同时卸载到多个服务车并行处理<sup>[30]</sup>。任务复制份数  $J$  对卸载可靠性的影响如图 19 所示。基于真实道路的仿真结果显示, 适量任务复制可在拓扑时

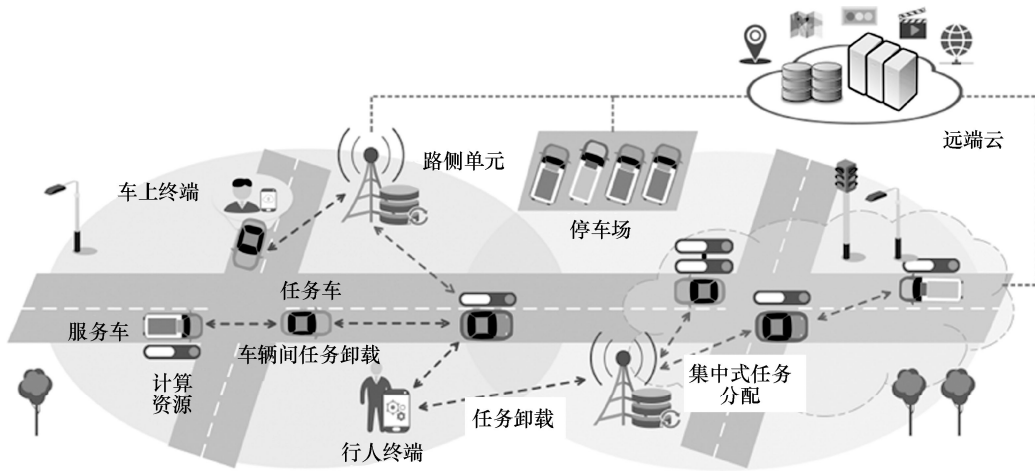


图 17 车载云计算架构

变、传输链路不稳定、车辆计算能力时变的环境下，使任务完成成功率达到 99.6%。进一步地，编码计算可使用较少的通信和计算资源获得良好的分集增益<sup>[27,31]</sup>，从而赋能移动智能体间的高可靠、低时延、高效协同。

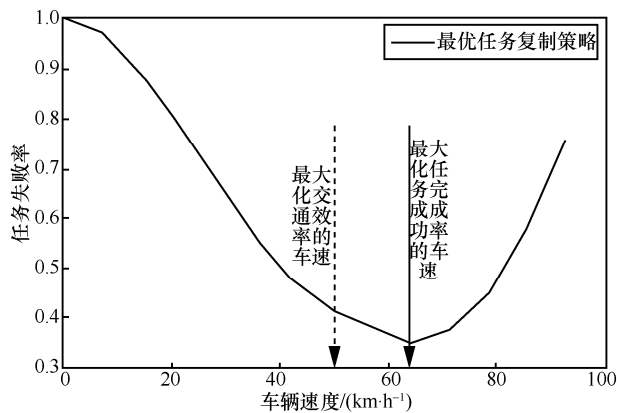


图 18 车辆速度与任务完成比例的关系

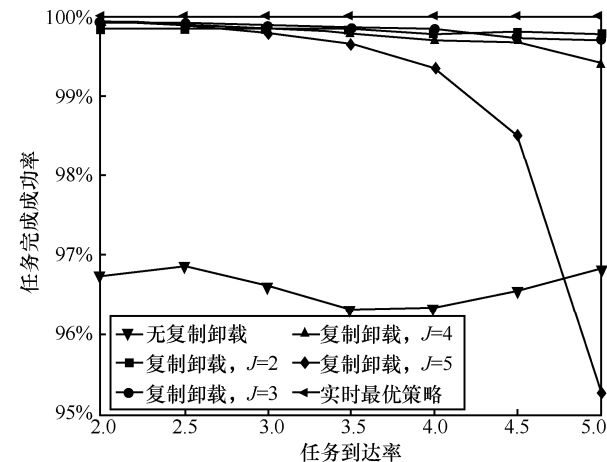


图 19 任务复制份数对卸载可靠性的影响

### 4 结束语

本文面向国家“双碳”战略需求，针对无线通信与边缘计算业务量高速增长所面临的能耗瓶颈，研究并总结出可在提高移动通信系统容量的同时大幅降低网络能耗的理论与方法。具体地，通过引入超蜂窝网络架构实现网络的柔性覆盖与弹性接入，使业务基站和边缘服务器在业务量较低时可以进入休眠状态，减少能量的浪费（即“节流”）。同时，大量引入可再生绿色能源（即“开源”），通过能量流与信息流的智能适配，大幅降低电网的能耗。进一步地，通过网络功能虚拟化、通信与计算资源的高效协同，以及移动智能体的分布式计算与协同等手段，实现绿色计算与人工智能算法。典型的解决方案展示了上述理论与方法的可行性与有效性，可支撑未来 5G 和 6G 网络作为智能化社会基础设施的可持续性演进。

### 参考文献：

[1] NIU Z S, TANGO: traffic-aware network planning and green operation[J]. IEEE Wireless Communications, 2011, 18(5): 25-29.  
 [2] 牛志升, 郑福春, 杨晨阳, 等. 基于超蜂窝架构的绿色通信专刊编者按[J]. 中国科学:信息科学, 2017, 47(5): 527-528.  
 NIU Z S, ZHENG F C, YANG C Y, et al. Editorial of the special issue on hyper-cellular communications[J]. SCIENTIA SINICA Informationis, 2017,47(5): 527-528.  
 [3] GUO X Y, NIU Z S, ZHOU S, et al. Delay-constrained energy-optimal base station sleeping control[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(5): 1073-1085.  
 [4] WU J, ZHOU S, NIU Z S. Traffic-aware base station sleeping control and power matching for energy-delay tradeoffs in green cellular networks[J]. IEEE Transactions on Wireless Communications, 2013, 12(8): 4196-4209.

- [5] GONG J, ZHOU S, NIU Z S. A dynamic programming approach for base station sleeping in cellular networks[J]. IEICE Transactions on Communications, 2012, E95-B(2): 551-562.
- [6] ZHANG S, GONG J, ZHOU S, et al. How many small cells can be turned off via vertical offloading under a separation architecture? [J]. IEEE Transactions on Wireless Communications, 2015, 14(10): 5440-5453.
- [7] CAO D X, ZHOU S, NIU Z S. Optimal combination of base station densities for energy-efficient two-tier heterogeneous cellular networks[J]. IEEE Transactions on Wireless Communications, 2013, 12(9): 4350-4362.
- [8] KAMITSOS I, ANDREW L, KIM H, et al. Optimal sleep patterns for serving delay-tolerant jobs[C]//Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking - e-Energy '10. New York: ACM Press, 2010: 31-40.
- [9] LENG B J, GUO X Y, ZHENG X, et al. A wait-and-see two-threshold optimal sleeping policy for a single server with bursty traffic[J]. IEEE Transactions on Green Communications and Networking, 2017, 1(4): 528-540.
- [10] JIANG Z Y, KRISHNAMACHARI B, ZHOU S, et al. Optimal sleeping mechanism for multiple servers with MMPP-based bursty traffic arrival[J]. IEEE Wireless Communications Letters, 2018, 7(3): 436-439.
- [11] LIU J C, KRISHNAMACHARI B, ZHOU S, et al. DeepNap: data-driven base station sleeping operations through deep reinforcement learning[J]. IEEE Internet of Things Journal, 2018, 5(6): 4273-4282.
- [12] ZHOU S, CHEN T J, CHEN W, et al. Outage minimization for a fading wireless link with energy harvesting transmitter and receiver[J]. IEEE Journal on Selected Areas in Communications, 2015, 33(3): 496-511.
- [13] HU C S, GONG J, WANG X L, et al. Optimal green energy utilization in MIMO systems with hybrid energy supplies[J]. IEEE Transactions on Vehicular Technology, 2015, 64(8): 3675-3688.
- [14] GONG J, THOMPSON J S, ZHOU S, et al. Base Station sleeping and resource allocation in renewable energy powered cellular networks[J]. IEEE Transactions on Communications, 2014, 62(11): 3801-3813.
- [15] ZHOU S, GONG J, ZHOU Z Y, et al. GreenDelivery: proactive content caching and push with energy-harvesting-based small cells[J]. IEEE Communications Magazine, 2015, 53(4): 142-149.
- [16] GONG J, ZHOU S, ZHOU Z Y, et al. Policy optimization for content push via energy harvesting small cells in heterogeneous networks[J]. IEEE Transactions on Wireless Communications, 2017, 16(2): 717-729.
- [17] 张平, 陶运铮, 张治. 5G 若干关键技术评述[J]. 通信学报, 2016, 37(7): 15-29.  
ZHANG P, TAO Y Z, ZHANG Z. Survey of several key technologies for 5G[J]. Journal on Communications, 2016, 37(7): 15-29.
- [18] ZHOU S, ZHAO T, NIU Z S, et al. Software-defined hyper-cellular architecture for green and elastic wireless access[J]. IEEE Communications Magazine, 2016, 54(1): 12-19.
- [19] LIU J C, ZHOU S, GONG J, et al. Statistical multiplexing gain analysis of heterogeneous virtual base station pools in cloud radio access networks[J]. IEEE Transactions on Wireless Communications, 2016, 15(8): 5681-5694.
- [20] LIU J C, XU S G, ZHOU S, et al. Redesigning fronthaul for next-generation networks: beyond baseband samples and point-to-point links[J]. IEEE Wireless Communications, 2015, 22(5): 90-97.
- [21] SHI W Q, HOU Y Z, ZHOU S, et al. Improving device-edge cooperative inference of deep learning via 2-step pruning[C]//Proceedings of IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops. Piscataway: IEEE Press, 2019: 1-6.
- [22] SHI W Q, ZHOU S, NIU Z S, et al. Joint device scheduling and resource allocation for latency constrained wireless federated learning[J]. IEEE Transactions on Wireless Communications, 2021, 20(1): 453-467.
- [23] SUN Y X, ZHOU S, NIU Z S, et al. Dynamic scheduling for over-the-air federated edge learning with energy constraints[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(1): 227-242.
- [24] WANG G C, ZHOU S, ZHANG S, et al. SFC-based service provisioning for reconfigurable space-air-ground integrated networks[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(7): 1478-1489.
- [25] 黄永明, 郑冲, 张征明, 等. 大规模无线通信网络移动边缘计算和缓存研究[J]. 通信学报, 2021, 42(4): 44-61.  
HUANG Y M, ZHENG C, ZHANG Z M, et al. Research on mobile edge computing and caching in massive wireless communication network[J]. Journal on Communications, 2021, 42(4): 44-61.
- [26] SUN Y X, ZHOU S, XU J. EMM: energy-aware mobility management for mobile edge computing in ultra dense networks[J]. IEEE Journal on Selected Areas in Communications, 2017, 35(11): 2637-2646.
- [27] ZHOU S, SUN Y X, JIANG Z Y, et al. Exploiting moving intelligence: delay-optimized computation offloading in vehicular fog networks[J]. IEEE Communications Magazine, 2019, 57(5): 49-55.
- [28] JIANG Z Y, ZHOU S, GUO X Y, et al. Task replication for deadline-constrained vehicular cloud computing: optimal policy, performance analysis, and implications on road traffic[J]. IEEE Internet of Things Journal, 2018, 5(1): 93-107.
- [29] SUN Y X, GUO X Y, SONG J H, et al. Adaptive learning-based task offloading for vehicular edge computing systems[J]. IEEE Transactions on Vehicular Technology, 2019, 68(4): 3061-3074.
- [30] SUN Y X, ZHOU S, NIU Z S. Distributed task replication for vehicular edge computing: performance analysis and learning-based algorithm[J]. IEEE Transactions on Wireless Communications, 2021, 20(2): 1138-1151.
- [31] ZHANG F, SUN Y X, ZHOU S. Coded computation over heterogeneous workers with random task arrivals[J]. IEEE Communications Letters, 2021, 25(7): 2338-2342.

#### [作者简介]



牛志升 (1964- ), 男, 河北秦皇岛人, 博士, 清华大学教授, 主要研究方向为通信话务理论、排队论、通信网络的流量控制与性能分析、无线网络的资源分配及跨层优化、通信与广播融合网络、绿色通信与网络等。

周盛 (1983- ), 男, 上海人, 博士, 清华大学副教授、博士生导师, 主要研究方向为绿色无线通信、车联网、移动云计算等。

孙宇璇 (1993- ), 女, 辽宁大连人, 博士, 清华大学在站博士后, 主要研究方向为移动边缘计算、边缘学习等。