



专题：通信与 AI 融合

AI 使能的 5G 节能技术

徐丹, 曾宇, 孟维业, 李力卡

(中国电信股份有限公司研究院 AI 研发中心, 北京 102209)

摘 要: 随着 5G 商用的全面开展, 5G 无线站点数目急剧增加, 5G 核心网需分层部署在区域/省/地市数据中心, 以及数据中心规模化发展, 导致能耗问题日益凸显。基于全网能耗主要占比, 调研 5G 接入网络、核心网络和数据中心的能源效率评估方法。介绍了 AI 使能的基站节能技术及试点应用方案、AI 应用于 5G 核心网的节能方式、AI 使能的数据中心节能技术和试点应用方案, 探讨了节能技术的挑战和未来的研究方向。对整体通信系统节能技术的总结和展望, 有助于提高对能源效率和绿色网络发展的认识。

关键词: 5G; AI; 能源效率; 接入网络; 核心网络; 数据中心

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021104

AI-enabled 5G energy-saving technology

XU Dan, ZENG Yu, MENG Weiye, LI Lika

AI Research & Development Center, Research Institute of China Telecom Co., Ltd., Beijing 102209, China

Abstract: With the development of 5G commercialization, the number of 5G wireless base station sites has increased sharply, 5G core networks need to be deployed in regions/provinces/cities data centers, as well as the data centers were developing on a larger scale, the problem of energy consumption was becoming increasingly prominent. Based on the main proportion of energy consumption in the whole network, the energy efficiency evaluation methods of 5G access network, core network and data center was investigated. The AI-enabled energy-saving technology of base stations and pilot application scheme, 5G core network energy-saving methods by using AI, as well as the AI-enabled energy-saving technology of data centers and pilot application scheme were introduced, and the challenges and future research directions of energy-saving technology were finally discussed. The summary and prospect on the energy-saving technology of the overall communication system, will help to improve the understanding of energy efficiency and green network development.

Key words: 5G, AI, energy efficiency, access network, core network, data center

收稿日期: 2021-04-10; 修回日期: 2021-05-10

基金项目: 国家重点研发计划基金资助项目 (No.2019YFB1802501, No.2019YFB1802504)

Foundation Items: The National Key Research and Development Program of China (No.2019YFB1802501, No.2019YFB1802504)

1 引言

5G 时代的到来将为海量设备提供网络服务,以满足泛在的连接及海量数据传输需求。根据《5G 时代运营商数据和存储架构白皮书》^[1]的预测,到 2025 年,5G 网络中每月数据量将达到 17.75 ZB,5G 网络容量需求显著提升。为了满足 5G 网络的容量需求,在当前的 5G 架构下,网络使用了更宽的频谱资源^[2],同时引入了大规模多输入多输出(massive multiple input multiple output, massive MIMO)技术,使得 5G 网络的能耗大幅增加^[3]。未来 5G 全面普及之后,预计到 2030 年,我国 5G 基站数量将达到 1 500 万个^[4],而单个 5G 基站的功耗是 4G 基站的 3~4 倍^[5],这会带来巨大的能源开销。依靠现有网络范式和体系结构来实现该目标将不可避免地导致能源紧缩,并带来严重的经济和环境问题。为应对全球能源短缺和环境保护面临的挑战,5G 系统的节能技术成为当前学术界和运营商研究的重点^[6]。

近年来,学术界已经开展了一系列研究,从理论上探索 5G 网络潜在的节能范式和体系架构。传统的网络节能主要通过合理的网络规划和多层关闭实现^[7],合理的网络规划意味着通过网络规划和优化,减少站点部署数量,以满足网络容量和节电的要求,提高覆盖效率,同时也可以减少无效的系统开销。然而 5G 网络的复杂性使得传统依赖于人工的运维手段无法满足节能需求,具体表现在:第一,解决方案不灵活。为了简化部署的复杂性,将在一个区域甚至整个网络中配置统一参数,而无须进行场景识别,这可能无法与实际流量变化相匹配;第二,该解决方案不支持可持续发展。一旦网络架构发生变化(例如小区数量的增加或减少、相邻小区关系的变化或流量的变化),先前的节能策略可能会完全无效;第三,该解决方案的适应性较差,需要大量的现场或远

程支持人员。随着人工智能(artificial intelligence, AI)技术的蓬勃发展,基于 AI 的智能化节能方案成为未来新的技术研究热点,通过网络预测、策略调整和优化以及实时 KPI 监视,基于 AI 设计闭环省电解决方案,在省电和网络性能之间找到平衡^[8]。

同时,运营商及产业界对 5G 节能也展开了一系列的实践探索,利用 AI 技术对网络历史数据进行学习,构建业务的模型,同时引入实时数据不断训练修正模型,实现对业务场景的精准预测,进而在当前环境下进行节能场景的识别、负荷的智能预测、节能策略的智能推荐,提高节能触发的灵活性,达到智慧节能的效果^[9]。

随着 5G 的增长,预计到 2025 年,无线接入网能耗占网络总能耗的 50.6%,数据中心占 23.3%,边缘和核心网占 2.2%^[10]。因此,在 5G 各个单元(如基站、核心网元、数据中心机房等)应用智慧节能技术势在必行,并将期待实现 5G 端到端能效提高。

2 能效模型

2.1 接入网络能效

在不考虑部署场景的情况下,能源效率可以表示为^[11]:

$$E = \sum_i \alpha_i \frac{V_i}{E_i} \quad (1)$$

其中, i 表示第 i 个载荷, V_i 表示网络每秒的流负载(byte), E_i 表示处理流量负载所消耗的能量(J), α_i 为流量负载级别的权重。

根据式(1),能效的提升可以分为两个方面:第一,利用更宽的频谱资源、更多的天线、更高的调制阶数等来增加流负载;第二,降低处理有效负载所消耗的能量。

2.2 核心网络能效

参考文献[12]中提出一种潜在的 5G 核心网络能效 KPI,可表示为:



$$EE_{5GC,DV} = \frac{DV_{5GC}}{EC_{5GC}} \quad (2)$$

其中, DV 表示 5G 核心网络测量的数据量, EC 表示核心网络消耗的能量。DV 测量包括 5G 核心网络仅用户面测量, 或者用户面和控制面测量。

根据式(2), 能效的提升可以分为两个方面: 第一, 通过提高单位时间有效流量负载; 第二, 降低处理有效负载所消耗的能量。

2.3 数据中心能效

在数据中心中, 将常用的电源使用效率(PUE)作为能效衡量指标^[13]:

$$PUE = \frac{\text{总用电量}}{\text{IT负载用电量}} \quad (3)$$

其中, 总用电量包括 IT 负载能耗、冷却系统的能耗、场地和核心房间的照明、电力与配电损耗等。由于冷却和设备散热、设备配置、设备机房环境和大气条件之间的关联, 在运维达到一定的成熟度后, 人力或专家经验无法满足进一步降低能耗的要求。降低数据中心总能耗, 可以通过动态负载调节优化服务器能耗^[14], 以及降低水冷主机、冷却塔、热交换器和水泵等能耗调节优化服务器能耗。

3 AI 使能的节能技术

3.1 基站节能

5G 基站是 5G 网络的核心设备, 提供无线覆盖功能, 实现有线通信网络与无线终端之间的无线信号传输。基站的架构、形态直接影响 5G 网络如何部署。由于频率越高, 信号传播过程中的衰减也越大, 与 4G 网络相比, 5G 网络需要更高的基站数量和密度, 随着覆盖范围、容量和站点数量的增加, 基站的功耗占移动网络的一半以上, 甚至高达 80%。因此, 5G 基站的节能迫在眉睫。

基站的能量消耗可以定义为:

$$P_B = \begin{cases} N_C N_{se} (P_0 + \Delta_p P_{\max} \chi), & 0 < \chi \leq 1 \\ N_C N_{se} P_{\text{sleep}}, & \chi = 0 \end{cases} \quad (4)$$

其中, N_C 是组件的载波数, N_{se} 表示每个站点的扇区数, Δ_p 表示负荷与功耗的斜率关系, χ 为传输资源元素(RE)的数量与总 RE 的数量之比, P_0 为基站最小输出功率, $P_{\max}$ 为射频模块每个载波上的最大输出功率, P_{sleep} 为基站睡眠模式下的功率。从式(4)可以看出, $\chi = 0$ 状态下的基站能耗更小^[15]。

在 LTE 网络中, 基站节能技术应用主要通过人工设定的门限来决定是否需要开关。这种方法参数设置保守, 节能效果有限, 且所有的参数都是由基站统一设置的, 无法在复杂多变的环境下做到对网络环境自适应。

在 5G 网络中, 灵活的参考信号设计使基站的睡眠模式更加高效。通过设计灵活的参考信号, 在不需要基站组件时将其关闭, 使得基站在更长的时间内处于 $\chi = 0$ 的状态, 从而降低了能耗 E_i 。例如, 基站的大多数组件都不使用时, 可以通过关闭其大部分组件的能量来使其进入睡眠模式。尤其在小区多层覆盖时, 通过流量预测和决定微基站的睡眠状态, 并将流量迁移到宏基站, 实现非常低的覆盖损耗, 在保障区域 QoS 的同时提供可能的最佳能源效率^[16]。

(1) AI 使能的 5G 基站节能技术

人工智能技术可以令基站节能技术更加智能化, 实现全网运行效率最优、综合节能效率最优的节能目标^[17]。通过数据预测和实时监控, 基于 AI 的基站节能可以分为以下 3 个方向。

- 基于时间的节能: 网络负载和用户行为可以通过历史数据分析来预测, 如自回归综合移动平均 (autoregressive integrated moving average, ARIMA) 模型、长短期记忆 (long short-term memory, LSTM) 和 Facebook 的 Prophet 用于时间序列预测^[18]。基于单元的时间窗口无须设置统一的参数, 而是可以确保更长的激活时间以节省功耗。
- 基于阈值的节能: 通过预测 RRC 连接的用户和 PRB 的使用情况, 将不同的触发阈值

用于不同的情况^[19]。将预测值与基于 15 min 粒度的实时流量进行比较, 预测准确性超过 90%, 这意味着逐个单元的策略和方案识别可提高可执行性。

- 基于 KPI 的节能: 实时监控网络 KPI (例如呼叫建立成功率、掉线率、切换成功率和吞吐量、平均功耗和用户体验) 以优化策略, 在 AI 省电中允许实时策略回滚^[20]。

AI 使能的基站节能方案需要采集无线端负载、感知数据等^[21], 使用 AI 技术识别业务场景^[22], 如商业区、居民区、高铁线路等。基于历史业务负荷数据训练预测模型, 在预测未来业务量增大时, 适当关闭一些节能手段, 或在预测的节能小区和节能时段触发该小区的负荷转移到其邻区, 负荷转移完成后立即触发该小区的节能^[23]。业务流量预测模型主要根据历史小区信息以及环境因素的特征, 利用神经网络找到这些特征值和模型的对应关系, 确定权重矩阵和梯度后构造, 表 1 给出一些网络流量预测模型。根据预测结果进行节能策略最优选择、实时调整、自动更新和执行, 实行差异化的阈值调整门限, 满足预设的 KPI 和节能目标^[24]。多基站协同节能的方法是将相邻基站的历史信息加入上述的模型进行联合模型的预测, 使预测结果更精确^[25]。同时也可以自动识别多层覆盖小区、基础覆盖小区和补热小区, 热点

覆盖小区可根据业务量变化、多网覆盖状态等进入节能模式, 实现节能效果。

(2) AI 使能的基站节能试点应用

某试点区域 5G 站点采用 AI 节能技术, 首先进行数据 (包括 OMC 数据、DPI 数据、MR/工参、话单数据等) 采集, 然后对数据进行融合处理包括清洗 (去噪、填充)、关联、稽核。通过试点小区的历史数据构建无监督学习分析潮汐结构、模式, 挖掘试点小区在不同周期 (月内、周内) 的无线资源、业务负荷等特征, 为其建立不同周期级别的潮汐档案。基于试点小区的潮汐结构和业务指标的时序预测结果, 结合场景识别、共覆盖分析情况, 生成试点小区的安全节能建议 (包含小区最佳节能时段、节能功能 (符号关断、载波关断、通道关断等) 的推荐), 然后配置 OMC 及指令集, 并自动完成执行策略实施后的网络性能、节能情况分析, 利用强化学习 (Q-learning) 与操作交互结果获得最大奖赏/惩罚 (预先设定) 优化参数调整。

5G 基站应用 AI 节能策略的同时亦评估对网络业务的影响, 支持突发事件早期感知智能刹车/唤醒。基于 AI 的预见性基站节能能够实现节能时段的精细选择, 延长节能时长, 同时保障用户感知。

3.2 5G 核心网络节能

核心网作为“管理中枢”, 负责管理数据,

表 1 网络流量预测模型

模型名称	模型特点描述	参考文献
改进的 LSTM 模型	考虑不同基站的流量之间的空间依赖性, 改进 LSTM 的结构	参考文献[26]
基于树的深度模型	基于树的模型使用了卷积层, 可以有效捕获流量的空间依赖信息	参考文献[27]
混合的时空预测深度学习模型	时间依赖性被 LSTM 捕获, 而空间依赖性被自动编码器封装, 用于建模来自相邻基站的历史流量信息, 从而将流量的空间依赖性纳入预测	参考文献[28]
流量预测架构	通过利用密集连接的中枢神经网络来捕获时空依赖关系	参考文献[29]
基于卷积神经网络的预测模型	通过建模时间和长距离的空间依赖关系, 该模型遵循了一个编码器对应一个解码器的范式, 组合了 LSTM 和 CNN 元素	参考文献[30]
基于图神经网络 (graph neural network, GNN)	利用图表示对网络流量时空依赖进行建模。将总数据流量分解为基站内流量和基站间流量, 这分别对应于基站覆盖范围内的用户服务流量和在不同建站覆盖区域之间移动的用户服务流量	参考文献[31]



对数据进行处理和路由交换。由于数据密集型应用程序不断普及，连接到互联网的设备数量不断增加，互联网流量已经呈指数级增长。在经济、环境和社会影响的驱动下，学术界和工业界对核心网的能耗开展了一系列研究，目标为有效提高能源效率。

随着软件定义网络（software defined network, SDN）和网络功能虚拟化（network functions virtualization, NFV）的出现，核心网元的功能实现可以通过 SDN 进行定义，进而通过网络云中的虚拟网络功能（virtualized network functions, VNF）来实现。虚拟化可能成为一种非常有用的降低能源消耗的方法，硬件充当同时运行多个应用程序的承载，可节省额外部署专用硬件和软件组件^[32]的成本。参考文献[33]提出了在接入网络和核心网络中使用虚拟网元节点，相比于使用物理网元节点执行单一功能的结构，虚拟网元有显著节能优势（如资源分配的灵活性、单个节点运行多个虚拟化接入/核心网功能）。虚拟化网元在不工作时休眠，根据工作负载随时准备激活这些功能。

在 3GPP 框架下，3GPP SA5 正在针对 5G 核心网以及切片的能效 KPI 进行标准化研究，利用 3GPP R17 的管理服务框架有可能实现利用管理服务对网元的能效 KPI 进行标准化测量和收集^[25]。

5G 核心网中应用 AI 技术，可以根据网络流量的变化趋势，调节虚拟网元资源占用情况，在不影响网络提供服务关键性能指标的前提下，降低能量消耗，提高网络的能量效率，从而达到节能的目的。

针对核心网控制面和用户面的资源使用和流量变化趋势调整资源分配，需要采集相关数据，包括但不限于：各网元实例的业务量、最大服务用户数、能耗数据、资源使用数据、NF 级性能测量数据和网络级 KPI 数据等。

基于网络流量的历史数据，需要对网络流量

变化趋势进行预测。网元实例流量的精准预测可以帮助推动节能的网络决策，然而网络流量在时间和空间上具有不均匀性、用户在一天的不同时刻和不同的地方有不同的 QoS 要求以及用户的移动性等使网络流量预测面临挑战，所以模型算法需要根据监测数据不断更新迭代。

基于网络流量的预测数据，可以智能地调整 VNF 的资源分配，例如根据 VNF 实例上的负载统计及预测情况，将资源利用率低的 VNF 实例上的负载迁移到特定的少数专用 VNF 实例上，从而产生更多的空负载的 VNF 实例，对这些 VNF 实例占用的资源进行重分配（或回收），提高核心网的能量效率 KPI。参考文献[34]提出一种基于半监督机器学习的资源需求新模型，参考文献[35]使用深度学习实现 NFV 中的资源分配。

以节能为目的，同时考虑用户的 QoS（QoE）需求，在满足用户需求的前提下，降低 VNF 的资源占用，从而在节能和网络需求层面达到均衡。参考文献[36]使用了 Facebook 数据中心的真实生活数据，在已经建立的包括移动管理实体、服务网关、分组数据网网关和策略、收费功能等基础设施之上进行了真实的估算，在工作队列下通过改变不同的因素（包括 VNF 的总数、可用的服务器总数以及传入的速度、整个系统容量和虚拟机的设置时间等不同参数），比较了消耗功率和服务器的响应时间，旨在基于动态自动缩放算法（dynamic auto scaling algorithm, DASA），为运营商提供一种更好地理解其成本与性能之间的权衡的方法。

3.3 数据中心节能

随着 5G 时代的到来，各类数据的需求增长，数据中心大量能耗带来的问题也引起了广泛的关注。

影响数据中心能耗主要由 3 个部分组成，分别是 IT 设备、制冷系统和配供电系统。其中，IT 设备和制冷系统能耗占数据中心总能耗的 80%以

上, 因此对 IT 设备和制冷系统的运行模式和控制方式进行优化对数据中心节能具有重要意义。

(1) AI 使能的数据中心节能技术

服务器主要进行数据传输和路由交换, 服务器设计通常会提供对数据的最大处理和路由交换能力, 但是在使用的数据量较少时, 服务所提供流量会产生冗余, 并且传输大部分时间远高于实际的使用流量, 也就是说即使服务器空闲也会发生大量的能耗。

通过使用人工智能技术, 可以基于历史数据训练学习得到业务模式变化和服务器运转对能耗的影响规律, 对未来业务负载进行预测并迭代更新, 当业务负载较少时进行业务迁移, 将一部分服务器转到休眠状态。此外, 当预测业务负载显著增加时, 可以将必要数量的服务器唤醒并重新进行业务负载的引导和分配。通过分析数据中心不同业务的负载特征, 发掘节能潜质, 实现设备休眠和唤醒, 达到绿色节能的效果^[37]。

数据中心通常采用过度冷却的方法, 将温度设定为较低的温度以确保服务器稳定运行不出现较热的问题, 这会导致极大的能源浪费。传统的温度控制方法依靠定时采集温度数据作为判定依据, 具有很大的时延。同时温度特性及环境的复杂性导致传统控制难以达到需求。

利用 AI 的机房节能策略可以以机房冷却系统为监控对象, 进行机房温度信息采集、温度数据采集、温度的预测和机房温度的控制^[38]。

数据采集除了机房温度信息, 还包括但不限于 IT 负载耗电量、末端空调耗电量、风机输出、温度设置、水冷机组各部件耗电量(此为水冷数据中心需要)等。

基于 AI 的智能控制通过对知识信息的学习和推理, 并通过启发式的方法求解, 神经网络控制是指在控制系统中应用神经网络技术, 对难以精确建模的复杂非线性对象进行神经网络模型辨识。神经网络控制的基本思想: 假设被控对象输

入 u 与输出 y 满足非线性关系 $y=g(u)$, 控制的作用就是确定最佳的输入, 从而使系统输出达到期望值。参考文献[39]利用 BP 算法作为控制器用于车站温度的控制; 参考文献[40]将径向神经网络和常规的 PID 结合, 达到控制室内温度的目的; 参考文献[41]将机房温度值和设备负载输入神经网络, 输出空调的节能控制参数, 使环境温度维持在正常范围。

数据中心利用 AI 技术对机房温度进行预测可以有效地解决利用经验对温度的预测和估计带来的可靠性低的问题^[42]。

参考文献[43]提出一种基于强化学习算法的数据中心机房控制方法, 获取当前时刻来自环境的状态量(包括机房各个通道中传感器的数值), 基于预先训练的强化学习模型得到当前时刻的控制参数, 控制数据中心机房的温度调节设备的工作状态, 直至数据中心机房的温度稳定, 达到节能的目的。

除了对末端机房制冷设备控制, 水冷数据中心的水冷机组的能效控制尤为重要。参考文献[44]利用神经网络模型建立 PUE 预测模型, 实施制冷系统参数设定, 优化了 PUE 指标。通过分析不同数据中心的 PUE 差异, 水冷机组可以采用控制制冷主机运行台数、降低冷却塔出水温度和提升冷冻水水温等措施实现能效优化^[45]。

(2) AI 使能的制冷设备节能试点应用

某试点数据中心采用冰蓄冷技术, 即可在用电低谷期使用电制冷机制冰, 将冷量储存在冰池中, 在用电高峰期将储存的冷量释放, 在不同季节不同负载的情况下可选择不同的工况运行模式。

在该试点数据中心的末端机房部署有机柜级温/湿度传感器, 采集 IT 设备和空调设备能耗的智能电表。通过采集 4 个月的历史数据, 采集频率为每小时 1 条, 数据包括机柜温度、末端空调运行参数、机柜功率、末端空调功率、主机功率、



主机运行参数等，训练机器学习模型，输出制冷设备的参数设置包括末端空调参数设置（空调启停、空调温湿度设置等）以及主机参数设置（主机运行模式、泵频率设定、冷凝温度设定值等），从而制定该试点数据中心制冷设备的节能调整策略并实施。通过应用 AI 技术建模分析制冷冗余，降低了试点机房的制冷能耗。

4 未来方向

虽然 AI 使能的节能技术在学术界引起了广泛的研究，并在产业界获得一些应用基础，然而在系统级建模、综合能效评估等方面仍存在一些挑战，基于此提出以下未来的研究和发展方向。

4.1 基于数字孪生的系统建模

数字孪生是虚拟世界与真实世界的动态连接，通过建立数字孪生模型，可实现物理系统与虚拟模型结合，具有可见性、预测性、假设分析支持和多系统互联的特征。

通过数字孪生建模接入网络，模型能够拟合和反映物理网络的环境和特点，可模拟虚拟系统在覆盖小区的用户移动，业务流量的增加、减小，以及不同节能功能使用时的响应，在复杂动态空间通过智能分析决策找到不同环境下最适宜的节能方式，既能够提高无线系统能效又不影响用户体验。

通过数字孪生建模数据中心，是一种实现数据中心精密空调节能运行的潜在方案。在虚拟模型中，可模拟 IT 负载、外部环境温度以及空调控制设定值等变更即时响应机房内的热环境变化，通过调试可以弥补经验数据的不足，输出空调控制的优化设定值，既能满足机房内的制冷需求，又能降低机房 PUE 指标。

数字孪生和人工智能结合应用在系统节能时，可模拟环境状态动态变化以及节能策略执行对系统的影响，有效保障在应用节能策略时避免影响系统性能。

4.2 系统级能效设计

绿色通信的研究必须从整体上考虑，这涵盖了通信生态系统的方方面面，有必要从系统的角度考虑能效解决方案。首先，除了系统使用阶段所消耗的能量外，用于制造电信设备的隐含能量在任何能效解决方案的评估中也起着重要的作用。任何用于能效改进的新体系结构和设备的引入都应该考虑系统生命周期中消耗的总能量^[46]。此外，系统某一点的能效优化可能导致其他点的次优结果，需采用端到端的视角来评估网络能源效率和能源消耗^[47]。

各类节能技术的有效集成有望优化系统能源消耗，但是整个系统的能源消耗建模、能效指标的使用和系统级能效设计面临挑战。系统级能效设计的目标是改善整个系统/网络的能效，同时不影响 QoS。

4.3 基于 AI 应用的数据约减

现有的机器学习（ML）算法能识别大型数据集的分析规律性，但是当在应用程序级关联时，其功效性和外部有效性不足^[48]。

当前 AI 算法应用在系统节能时，依赖于大量数据。在存量数据中心中，数据条件具备情况差异很大，按照具备较好数据的数据中心训练的 AI 模型可能在其他机房完全不具备部署条件，因为很多数据中心在设计建设之初未考虑各类机房状态和参数监控。数据中心应用 AI 技术需要基础的训练数据集，但是对于已投入使用的数据中心，各类数据采集涉及的改造是一个很大的挑战。

为了使 AI 算法具备更好的部署性和外部适用性，需要研究训练数据缩减方法以开发完全自主的系统。

4.4 绿色人工智能

专用硬件平台不断增长的计算能力为 ML 解决方案提供了基础，然而不断提高模型精度所需的计算资源和产生的能源消耗不容忽视。利用分布式、计算高效和能量高效的方式训练 ML 将成

为一种研究方向,如联邦学习^[49],可缓解传统的、集中的 ML 方法造成的低能效问题。此外,当找到接近最优的解决方案时,也可以使用定制的及时停止来终止训练过程,从而减少训练 ML 模型所需的迭代次数和相关的能量消耗^[50]。

ML 的绿色应用还需要在优化算法的设计阶段评估计算和能源消耗,即在评估 ML 模型时除了准确性和优化指标外,还必须考虑计算效率和能源消耗。使用这些指标,可以设计得到收敛速度更快和消耗计算资源和能耗更少的 ML 模型。

5 结束语

对 5G 接入网络、核心网络和数据中心的能源效率评估方法进行了调研,基于降低系统能耗及提高能效,介绍了 AI 使能的基站节能技术,包括无线网络数据采集、场景识别、业务流量预测、节能策略推荐、基站协同节能控制及试点应用;其次介绍了 AI 应用于 5G 核心网的节能方式,包括核心网控制面和用户面的数据采集、网络流量预测以及 VNF 的资源分配;并介绍了 AI 用于降低数据中心制冷设备能耗的技术和试点应用;最后提出了节能技术的未来研究和发展方向,提供了端到端 AI 节能技术的全景。“能效管理、绿色发展”是未来网络发展的目标,端到端 AI 智能化节能手段的应用能够缓解运营商的能耗压力,并提升运维效率,将促进企业转型升级及可持续发展,从而实现经济、社会、生态效益的多赢局面。

参考文献:

- [1] 华为技术有限公司. 5G 时代运营商数据和存储架构白皮书[R]. 2020.
Huawei Technologies Co., Ltd. White paper on operator data and storage architecture in 5G era [R]. 2020.
- [2] GEORGAKOPOULOS A, MARGARIS A, TSAGKARIS K, et al. Resource sharing in 5G contexts: achieving sustainability with energy and resource efficiency[J]. IEEE Vehicular Technology Magazine, 2016, 11(1): 40-49.
- [3] GAO H Y, SU Y M, ZHANG S B, et al. Antenna selection and power allocation design for 5G massive MIMO uplink networks[J]. China Communications, 2019, 16(4): 1-15.
- [4] 赛迪顾问. 5G 产业发展白皮书(2020) [R]. 2020.
CCID Consulting. 5G industry development white paper (2020)[R]. 2020.
- [5] 岑祺. 5G 基站市电建设及改造方案[J]. 信息通信, 2019, 32(12):210-213.
CEN Q. 5G base station municipal power construction and transformation plan[J]. Information & Communications, 2019, 32(12): 210-213.
- [6] WU Q Q, LI G Y, CHEN W, et al. An overview of sustainable green 5G networks[J]. IEEE Wireless Communications, 2017, 24(4): 72-80.
- [7] OH E, KRISHNAMACHARI B, LIU X, et al. Toward dynamic energy-efficient operation of cellular network infrastructure[J]. IEEE Communications Magazine, 2011, 49(6): 56-61.
- [8] MEMON M L, MAHESHWARI M K, SAXENA N, et al. Artificial intelligence-based discontinuous reception for energy saving in 5G networks[J]. Electronics, 2019, 8(7): 778.
- [9] LOPEZ-PEREZ D, DE DOMENICO A, PIOVESAN N, et al. A survey on 5G energy efficiency: massive MIMO, lean carrier design, sleep modes, and machine learning[J]. arXiv preprint arXiv: 2101.11246, 2021.
- [10] LORINCZ J, CAPONE A, WU J S. Greener, energy-efficient and sustainable networks: state-of-the-art and new trends[J]. Sensors (Basel, Switzerland), 2019, 19(22): E4864.
- [11] LI Y N R, CHEN M Z, XU J, et al. Power saving techniques for 5G and beyond[J]. IEEE Access, 2020(8): 108675-108690.
- [12] 3GPP. Study on new aspects of energy efficiency (EE) for 5G: TR 28.813 V0.5.0[S]. 2021.
- [13] ITU-T. Total network infrastructure energy efficiency metrics: L.1332[S]. 2018.
- [14] CHAO L, QOUNEH A, TAO L. iSwitch: coordinating and optimizing renewable energy powered server clusters[J]. ACM Sigarch Computer Architecture News, 2012, 40(3): 512-523.
- [15] DEBAILLIE B, DESSET C, LOUAGIE F. A flexible and future-proof power model for cellular base stations[C]// Proceedings of 2015 IEEE 81st Vehicular Technology Conference (VTC Spring). Piscataway: IEEE Press, 2015: 1-7.
- [16] DONEVSKI I, VALLERO G, MARSAN M A. Neural networks for cellular base station switching[C]//Proceedings of IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2019: 738-743.
- [17] 张志荣, 许晓航, 朱雪田, 等. 基于 AI 的 5G 基站节能技术研究[J]. 电子技术应用, 2019, 45(10): 1-4.
ZHANG Z R, XU X H, ZHU X T, et al. Research on energy saving technology of 5G base station based on AI[J]. Application of Electronic Technique, 2019, 45(10): 1-4.



- [18] LI R P, ZHAO Z F, ZHOU X, et al. Intelligent 5G: when cellular networks meet artificial intelligence[J]. IEEE Wireless Communications, 2017, 24(5): 175-183.
- [19] GAO Y, CHEN J J, LIU Z, et al. Machine learning based energy saving scheme in wireless access networks[C]//Proceedings of 2020 International Wireless Communications and Mobile Computing (IWCMC). Piscataway: IEEE Press, 2020: 1573-1578.
- [20] SOLIMAN S S, SONG B. Fifth generation (5G) cellular and the network for tomorrow: cognitive and cooperative approach for energy savings[J]. Journal of Network and Computer Applications, 2017, 85: 84-93.
- [21] SDN/NFV/AI 标准与产业推进委员会. 智能化基站节能研究[R]. 2020.
SDN/NFV/AI Standards and Industry Promotion Committee. Research on energy saving of intelligent base station [R]. 2020.
- [22] 张青. 基于 AI 的 5G 基站节能场景智能识别[C]//5G 网络创新研讨会 (2019) 论文集. 移动通信, 2019(Z2): 123-127.
ZHANG Q. Intelligent identification of energy saving scenario of 5G base station based on AI [C]// Symposium on 5G Network Innovation (2019). Mobile Communication, 2019(Z2): 123-127.
- [23] 任嘉鹏. 基于机器学习的流量预测及基站休眠方法研究[D]. 长春: 吉林大学, 2020.
REN J P. Research on machine learning based traffic prediction and base station sleep mode[D]. Changchun: Jilin University, 2020.
- [24] 李继蕊, 李小勇, 高云全, 等. 5G 网络下移动云计算节能措施研究[J]. 计算机学报, 2017, 40(7): 1491-1516.
LI J R, LI X Y, GAO Y Q, et al. Energy saving research on mobile cloud computing in 5G[J]. Chinese Journal of Computers, 2017, 40(7): 1491-1516.
- [25] 易芝玲, 孙奇, 吴杰, 等. 人工智能在 5G 无线网络中的标准与应用进展[J]. 信息通信技术与政策, 2020(9): 23-30.
CHIH-LIN I, SUN QI, WU JIE, et al. AI in 5G networks—usage scenario and standardization progress[J]. Information and Communications Technology and Policy, 2020(9): 23-30.
- [26] ZHANG X, YOU J L. A gated dilated causal convolution based encoder-decoder for network traffic forecasting[J]. IEEE Access, 2020(8): 6087-6097.
- [27] HOSSAIN M S, MUHAMMAD G. A deep-tree-model-based radio resource distribution for 5G networks[J]. IEEE Wireless Communications, 2020, 27(1): 62-67.
- [28] WANG J, TANG J, XU Z Y, et al. Spatiotemporal modeling and prediction in cellular networks: a big data enabled deep learning approach[C]//Proceedings of IEEE INFOCOM 2017 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2017: 1-9.
- [29] ZHANG C, ZHANG H, YUAN D, et al. Citywide cellular traffic prediction based on densely connected convolutional neural networks[J]. IEEE Communications Letters, 2018, 22(8): 1656-1659.
- [30] ZHANG C Y, PATRAS P. Long-term mobile traffic forecasting using deep spatio-temporal neural networks[C]//Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing. New York: ACM Press, 2018: 231-240.
- [31] WANG X, ZHOU Z M, XIAO F, et al. Spatio-temporal analysis and prediction of cellular traffic in metropolis[J]. IEEE Transactions on Mobile Computing, 2019, 18(9): 2190-2202.
- [32] AL-QUZWEENI A N, LAWEY A Q, ELGORASHI T E H, et al. Optimized energy aware 5G network function virtualization[J]. IEEE Access, 2019(7): 44939-44958.
- [33] AL-QUZWEENI A, EL-GORASHI T E H, NONDE L, et al. Energy efficient network function virtualization in 5G networks[C]//Proceedings of 2015 17th International Conference on Transparent Optical Networks (ICTON). Piscataway: IEEE Press, 2015: 1-4.
- [34] KIM H G, LEE D Y, JEONG S Y, et al. Machine learning-based method for prediction of virtual network function resource demands[C]//Proceedings of 2019 IEEE Conference on Network Softwarization (NetSoft). Piscataway: IEEE Press, 2019: 405-413.
- [35] XU J, WANG J Y, QI Q, et al. IARA: an intelligent application-aware VNF for network resource allocation with deep learning[C]//Proceedings of 2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). Piscataway: IEEE Press, 2018: 1-3.
- [36] REN Y, PHUNG-DUC T, CHEN J C, et al. Dynamic auto scaling algorithm (DASA) for 5G mobile networks[C]//Proceedings of 2016 IEEE Global Communications Conference (GLOBECOM). Piscataway: IEEE Press, 2016: 1-6.
- [37] 徐丹, 王海宁. AI 助力 5G 发展研究[J]. 中国电信建设, 2019, 31(3): 46-54.
XU D, WANG H N. AI helps 5G development research[J]. China Telecommunications Construction, 2019, 31(3): 46-54.
- [38] 吴亚奇. 数据中心机房节能方法研究[D]. 苏州: 苏州科技大学, 2019.
WU Y Q. Research on energy saving method of data center computer room[D]. Suzhou: Suzhou University of Science and Technology, 2019.
- [39] 邹燕. IDC 机房温度调控设计研究[D]. 上海: 上海工程技术大学, 2016.
ZOU Y. The reserch on the control of IDC room temperature[D]. Shanghai: Shanghai University of Engineering Science, 2016.
- [40] 屈毅, 宁铎, 赖展翅, 等. 温室温度控制系统的神经网络 PID 控制[J]. 农业工程学报, 2011, 27(2): 307-311.
QU Y, NING D, LAI Z C, et al. Neural networks based on PID control for greenhouse temperature[J]. Transactions of the Chinese Society of Agricultural Engineering, 2011, 27(2): 307-311.
- [41] 张志鹏, 龚慧钦, 陈庆. 基于人工智能的机房空调控制方法、

- 装置及系统: CN110285532A[P]. 2019-09-27.
- ZHANG Z P, GONG H Q, CHEN Q. Machine room air conditioner control method, device and system based on artificial intelligence: CN110285532A[P]. 2019-09-27.
- [42] 徐慧姣. 数据中心机房的节能减排技术及发展[J]. 通讯世界, 2019, 26(3): 251-259.
- XU H J. Energy saving and emission reduction technology and development of data center machine room [J]. Telecom World, 2019, 26(3): 251-259.
- [43] 周鹏飞, 马亮. 一种基于人工智能优化机房能耗效率的方法及装置: CN109890176A[P]. 2019-06-14.
- ZHOU P F, MA L. Method and device for optimizing machine room energy consumption efficiency on the basis of artificial intelligence: CN109890176A[P]. 2019-06-14.
- [44] 杨震, 赵静洲, 林依挺, 等. 数据中心 PUE 能效优化的机器学习方法[J]. 系统工程理论与实践, 2020: 1-12.
- YANG Z, ZHAO J Z, LIN Y T, et al. Machine learning approach for energy efficiency optimization of PUE in data centers[J]. Systems Engineering-Theory & Practice, 2020: 1-12.
- [45] 姜洁良. 数据中心能效案例及节能运行[J]. 电信科学, 2019, 35(2): 95-104.
- LOU J L. Data center energy efficiency case and energy saving operation[J]. Telecommunications Science, 2019, 35(2): 95-104.
- [46] BUZZI S, I C L, KLEIN T E, et al. A survey of energy-efficient techniques for 5G networks and challenges ahead[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(4): 697-709.
- [47] CHEN T, YANG Y, ZHANG H G, et al. Network energy saving technologies for green wireless access networks[J]. IEEE Wireless Communications, 2011, 18(5): 30-38.
- [48] KAUR J, KHAN M A, IFTIKHAR M, et al. Machine learning techniques for 5G and beyond[J]. IEEE Access, 2021(9): 23472-23488.
- [49] KAIROUZ E B P, MCMAHAN H B. Advances and open problems in federated learning[J]. Foundations and Trends® in Machine Learning, 2021, 14(1). Doi:10.1561/22000000083.
- [50] PRECHELT L. Early stopping — but when?[M]//Lecture Notes

in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 53-67.

[作者简介]



徐丹 (1993-), 女, 中国电信股份有限公司研究院 AI 研发中心工程师, 主要研究方向为网络人工智能、网络切片、基站节能、机房节能。



曾宇 (1978-), 男, 博士, 中国电信股份有限公司研究院 AI 研发中心 IDC 节能产品线主管, 主要研究方向为 AI/5G、网络智能化。



孟维业 (1980-), 男, 中国电信股份有限公司研究院 AI 研发中心 IDC 节能架构总监, 主要研究方向为云原生、微服务及 AI 算法。



李力卡 (1975-), 男, 中国电信股份有限公司研究院 AI 研发中心教授级高级工程师, 长期从事网络智能运营技术研究、大数据及 AI 产品研发方面的工作, 目前负责基站智慧节能系统设计研发工作。