

基于DDPG的基站节能决策优化方法

杨忱逊, 高月红, 杨鸿文

北京邮电大学信息与通信工程学院, 北京

收稿日期: 2025年3月25日; 录用日期: 2025年4月18日; 发布日期: 2025年4月27日

摘要

在通信网络中, 当用户通信需求随时间降低时, 基站可执行两种节能决策: 一是直接关断低负载基站, 将用户的服务转移至其他基站; 二是通过功率分级调节逐级降低基站的能耗, 同时维持通信服务的稳定性。本研究基于深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)的基站节能决策优化方法, 将综合考虑用户通信服务质量和系统能耗, 选择性能更优的节能决策。实验结果表明, 相较于基站保持高功率运行的基线实验, 文章提出的两种节能决策能够使系统能耗和时延的综合性能分别提升47.5%和56.4%。对比两种节能决策, 直接关断表现出更快的收敛特性, 而功率分级调节则在能耗和时延的整体性能上更具优势, 比关断决策高出16.9%。

关键词

基站节能, 强化学习, 决策优化

Base Station Energy-Saving Decision Optimization Method Based on DDPG

Chenxun Yang, Yuehong Gao, Hongwen Yang

School of Information and Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing

Received: Mar. 25th, 2025; accepted: Apr. 18th, 2025; published: Apr. 27th, 2025

Abstract

In communication networks, when the user communication demand decreases over time, base stations can implement two energy-saving decisions: one involves directly shutting down low-load base stations and transferring user services to other operational base stations; the other entails gradually reducing the energy consumption of base stations through hierarchical power adjustment, ensuring that communication services remain stable. In this paper, the base station energy-

saving decision optimization method based on Deep Deterministic Policy Gradient (DDPG) will comprehensively consider the user communication service quality and system energy consumption and select the optimal energy-saving decision. The experimental results show that, compared with the baseline experiment in which base stations maintain high power operation, the two energy-saving decisions proposed in this paper can improve the comprehensive performance of system energy consumption and delay by 47.5% and 56.4%, respectively. When comparing the two energy-saving decisions, direct shutdown shows faster convergence, while hierarchical power adjustment has more advantages in the overall performance of energy consumption and delay, representing a 16.9% improvement over the shutdown decision.

Keywords

Base Station Energy Saving, Reinforcement Learning, Decision Optimization

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

通信网络中, 基站的设计与部署通常是基于满足用户峰值速率需求的目标, 因此基站往往按照最大网络负载值为用户提供服务[1]。然而, 现实场景中用户流量通常随时间动态分布, 导致大部分时间内, 部分基站会处于轻载或空载状态, 但设备仍持续运行, 使得能耗居高不下[2]。为减少不必要的能源消耗, 可以在保障用户通信质量并遵循基站最大发射功率约束的条件下, 执行合理的基站节能决策, 关断低负载基站, 并将其业务负载迁移至周边基站[3], 从而实现闲时能耗的有效控制和性能优化。

传统的基站节能决策是随机关断网络中的某些基站, [4] [5]或根据基站和用户间的距离来决定是否关断[6] [7]。文献[4]通过综合考虑网络平均覆盖概率和用户功耗来量化基站关断对网络性能的影响, 构建了一个基站能耗最小化问题, 确定了关断宏基站(Macro Base Station, MBS)的最优概率, 实现在降低能耗的同时维持通信覆盖率。文献[5]发现当基站深度睡眠和不连续传输(Discontinuous Transmission, DTX)同时应用于提升网络能效时, 基站深度睡眠会增加剩余活动基站的负载, 减少它们通过 DTX 节省的能源, 通过权衡深度睡眠节能和 DTX 节能, 提出同时应用两种机制时的最佳基站节能策略。文献[6]提出的启发式算法是基于用户的位置和其所需的数据速率, 按照距离由近及远地逐步激活基站, 在保证用户连接需求的同时, 显著降低基站的激活比例来节省能量。文献[7]提出, 与功耗随业务负载而显著增加的宏基站相比, 微基站的功耗相对平稳且受负载影响较小, 在用户均匀分布的情况下, 优先关断距离宏基站较近的微基站, 宏基站以较低的功耗为用户提供服务, 可以降低异构网络的整体能耗。

同时, 强化学习(Reinforcement Learning, RL)也已经越来越多地应用于基站节能决策。文献[8]面向具有宏基站和多个微基站的密集异构网络背景下的节能决策问题, 提出一种基于深度 Q-Learning 的算法, 在基站切换需要额外的能源消耗时显著节能。文献[9]考虑到环境的动态性质, 提出了一种基于深度强化学习的动态基站切换框架, 同时引入了模仿学习(Imitation Learning, IL)的思想, 通过学习专家的动作来更新策略, 能够有效降低功耗。

以上相关研究, 主要聚焦于基站执行某一种特定节能决策的效果, 并未讨论基站执行节能决策面临多种选择时的优化方式。为应对用户需求随时间降低带来的影响, 基站通常面临两种节能决策: 一种决策是直接关断低负载基站, 将用户的通信连接切换至其他基站。这种方式能够迅速降低基站的恒定能耗,

但用户切换至其他基站的过程中将带来更多的传输能耗，并且可能降低用户通信服务质量；另一种决策是通过功率分级调节来逐级降低基站的能耗，同时维持通信服务的稳定性。本文研究的基于深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG) [10]的基站节能决策优化方法将综合考虑通信服务质量和系统能耗，通过对比，选择执行性能更优的节能决策。

2. 系统模型

2.1. 研究场景建模

本文的研究场景设置为多小区中有 M 个基站和 N 名静止用户，初始状态为 1 名用户连接 1 个基站。假设其中 K 名用户的通信业务 b_i 逐渐减小， T 时间段后，通信业务 b_i 消失，其余 $N-K$ 名用户的通信业务需求保持不变。此时基站需满足的目标为降低系统能耗，同时保障通信连接稳定。若初始基站难以满足用户的通信业务需求，用户设备选择切换接入附近的基站。通信网络拓扑如图 1 所示。

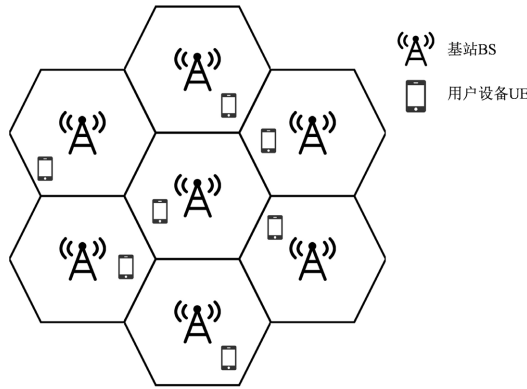


Figure 1. Communication network topology diagram

图 1. 通信网络拓扑图

2.2. 信道参数建模

本文参考了应用于 ITU 城市异构网络场景的路径损耗场景，选取大尺度衰落模型作为信道传播模型。信号传输中的基本路径损耗 PL_b 的计算公式如下：

$$PL_b = FSPL(d, f_c) + \varphi + \varepsilon \quad (1)$$

基本路径损耗 PL_b 由三部分组成， φ 是阴影衰落， ε 是天线增益， $FSPL$ 是自由空间路径损耗：

$$FSPL(d, f_c) = 32.45 + 20 \log_{10}(f_c) + 20 \log_{10}(d) \quad (2)$$

其中， f_c 为基站发射载波的频率， d 为用户与基站之间的直线距离。

2.3. 用户通信业务建模

在此研究场景下，每个基站的传输带宽固定，设基站 m 的带宽为 B_m 。基站按照服务区域内用户的通信业务需求向用户按需分配带宽。设基站 m 服务的用户数目为 Num_m ，并设基站 m 下用户 n 的通信业务需求为 $b_{m,n}$ ，那么基站 m 下的用户 n 所分配的带宽 $B_{m,n}$ 为：

$$B_{m,n} = \frac{B_m b_{m,n}}{\sum_{n=1}^{Num_m} b_{m,n}} \quad (3)$$

用户业务的传输速率 $V_{m,n}$ 根据香农公式计算可得:

$$V_{m,n} = B_{m,n} \log_2 \left(1 + \frac{p_{m,n} 10^{\frac{PL_b}{10}}}{\sigma^2} \right) \quad (4)$$

其中, $p_{m,n}$ 是用户设备的发射功率, σ^2 是信道噪声功率。根据此速率传输, 则用户与基站之间的传输时延 (Transmitting Delay) d^T 为:

$$d_{m,n}^T = \frac{b_{m,n}}{V_{m,n}} \quad (5)$$

用户与基站之间的传输能耗 (Transmitting Consumption) c^T 可通过用户设备的发射功率和传输时延相乘得到:

$$c_{m,n}^T = p_{m,n} \times d_{m,n}^T \quad (6)$$

3. 基于 DDPG 的基站节能决策优化方法

3.1. 基站节能问题建模

在本文的研究场景下, 用户的通信业务需求随时间逐渐降低。为避免所有基站因遵循传统的基站传输策略, 始终保持开启导致系统能耗过高, 基站可结合用户分布和业务负载情况执行智能节能决策, 在用户流量低负载时, 控制部分基站进入低功耗模式或关断状态, 完成降低系统能耗的目标。用户与基站之间的传输能耗为 c^T , 可由前文式(6)得到。用户从原连接基站切换至另一基站时的切换损耗 (Switch Consumption) 为 c^S , 如果相邻时刻用户 n 选择接入的基站不变, 则切换损耗 $c_{m,n}^S$ 等于 0, 反之:

$$c_{m,n}^S = \text{ABS}(\text{RSRP}_n^{\max}, \text{RSRP}_n) \quad (7)$$

其中, $\text{RSRP}_n^{\max}$ 为用户 n 当前时刻最佳的 RSRP, RSRP_n 为用户 n 上一时刻的 RSRP。基站在开启或关断状态下的恒定能耗 (Constant Consumption) 为 c^{Con} 。根据前文研究场景建模可得, K 个基站的功率等级随 K 名通信业务需求减小的用户 k 可调, 设 α 为功率等级系数, $\alpha \in (0, 1]$, 其余 $M-K$ 个基站保持满功率状态, 能耗为 $P_{\max}$ 。则系统中的基站恒定能耗 c^{Con} 可表示为:

$$c^{\text{Con}} = K \times P_{\max} \times \alpha + (M - K) \times P_{\max} \quad (8)$$

该研究场景下的优化目标为系统的总能耗降低, 且用户的通信服务质量提高, 即系统的总时延降低, 故系统的总体优化目标 G 可以表示如下:

$$\min G = \min \sum_{m=1}^M \sum_{n=1}^N (d_{m,n}^T + c_{m,n}^T + c_{m,n}^S) + c^{\text{Con}} \quad (9)$$

基站的节能决策优化被建模成一个对系统总体目标 $\min G$ 求最优化的问题。针对该问题, 本文设计了两种基站节能决策方案, 并研究了它们的优化方法, 旨在降低该研究场景下的系统能耗和时延。

3.2. 传统基站固定功率方案

传统基站不具备自主决策的能力, 基站优先保证现有用户的流量稳定性, 因此基站不会自主关断, 功率等级也不会随着用户需求降低, 直至用户通信业务需求降至 0, 基站始终保持满功率状态运行。以该传统基站固定功率方案作为基线实验。该方案下, 用户 n 选择接入的基站始终保持不变, 系统的切换损耗 c^S 等于 0, 系统中的基站恒定能耗 c^{Con} :

$$c^{Con} = M \times P_{\max} \quad (10)$$

3.3. 基站节能决策方案

为了解决系统总体目标 $\min G$ 的最优化问题，本节设计了两种基站节能决策，即基站面对本文提出的研究场景可产生两种节能决策。然而，基站针对同一时空的同一场景只能选择一种智能决策。

在第一种节能决策下，基站的决策被设定为“开启”和“关断”两个等级。基站根据其服务的用户业务量，当用户 k 通信需求降低时，业务量小的基站将直接进入关断状态，原本接入它的用户 k 将切换至附近处于开启状态的基站进行通信连接。在满足用户需求的条件下迅速关断基站，基站恒定能耗 c^{Con} 降低。但由于用户 k 需要切换距离较远的基站，因此系统中存在切换损耗 c^S ，且传输能耗 c^T 将上涨。由于基站只有“开启”和“关断”两种状态，因此式(8)中的功率等级系数 α 为定值，关断基站的最低能耗为 $P_{\min}$ 。则系统中的基站恒定能耗 c^{Con} ：

$$c^{Con} = K \times P_{\min} + (M - K) \times P_{\max} \quad (11)$$

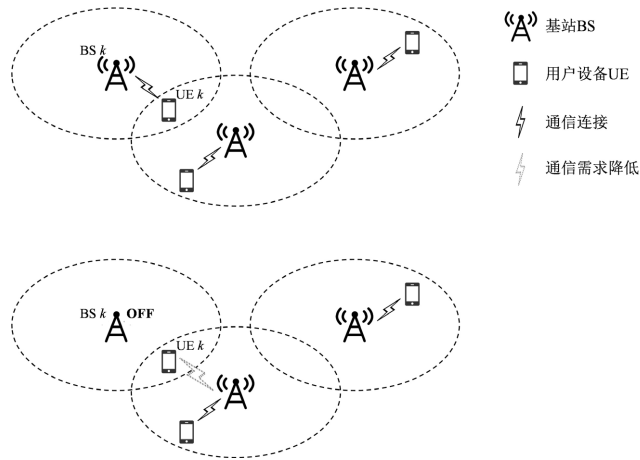


Figure 2. Energy-saving decision I: shutdown low-load base stations
图 2. 节能决策一：关断低负载基站

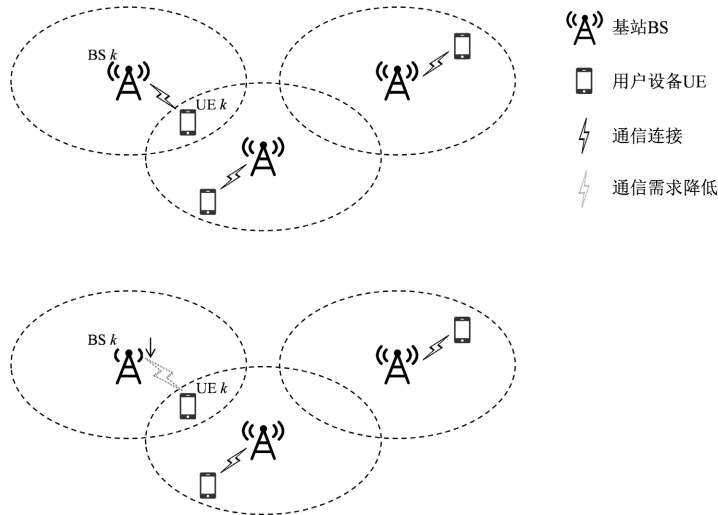


Figure 3. Energy-saving decision II: hierarchical power adjustment
图 3. 节能决策二：功率逐级调节

节能决策一(直接关断低负载基站)的情况如图 2 所示,其中用户 k 通信业务需求降低,初始状态下为用户 k 提供服务的基站 k 随即关断,用户 k 切换连接至就近基站。

在第二种节能决策下,基站的决策被设定为允许开启 N_p 个功率等级。基站支持根据用户业务量的变化情况进行不同的功率选择,实现更加精确的功率等级控制,同时避免用户 k 切换至其他基站进行连接,降低切换损耗 c^S 与传输能耗 c^T 。但由于没有直接关断基站,导致瞬时恒定能耗 c^{Con} 相较第一种节能决策偏高。该方案下,用户 n 选择接入的基站始终保持不变,因此系统的切换损耗 c^S 等于 0。系统中的基站恒定能耗 c^{Con} 可以直接用式(8)表示,随功率等级系数 α 变化。节能决策二(功率按等级调节)的情况如图 3 所示,其中用户 k 通信业务需求降低,基站 k 仍为用户提供通信连接,它的功率逐级降低。

3.4. 基于 DDPG 的基站节能决策优化方法

在强化学习任务中,智能体通过与环境的交互来学习最优的行动策略,达成奖励值的最大化。因此可以应用强化学习来实现系统能耗和用户通信服务质量总体目标 $\min G$ 的最优化。下面介绍应用于本问题的强化学习算法所设计的状态空间(State)、动作空间(Action)和奖励函数(Reward)。

1) 状态空间

部署在基站的智能体从基站处收集当前网络环境的状态信息,其中涵盖了用户位置分布、接入基站、业务需求、传输速率等用户状态,以及基站服务用户数量、信道参数等基站状态。选取用户传输速率和基站服务用户数量两项关键要素形成状态空间 S ,则该状态空间 S 可表示为如下集合:

$$S = \{s_1, s_2, \dots, s_m, \dots, s_M\} \quad (12)$$

其中, s_m 表示基站 m 当前的状态信息,它包括基站 m 服务的用户数量 s_{m0} 和基站 m 所服务用户的平均带宽 s_{m1} ,具体表示为:

$$s_m = \{s_{m0}, s_{m1}\} \quad (13)$$

$$s_{m0} = Num_m \quad (14)$$

$$s_{m1} = \frac{\sum_{n=1}^{Num_m} B_{m,n}}{Num_m} \quad (15)$$

2) 动作空间

智能体获取状态空间 S 后,经过强化学习训练,向基站输出动作空间 A ,执行当前状态下的节能决策,即立刻执行关断动作或按照功率等级参数逐级调节功率。节能决策一的动作空间 A_1 可表示为:

$$A_1 = \{a_1, a_2, \dots, a_m, \dots, a_M\}, a_m \in \alpha_1 \quad (16)$$

$$\alpha_1 = \{P_{\min}, P_{\max}\} \quad (17)$$

节能决策二的动作空间 A_2 可表示为:

$$A_2 = \{a_1, a_2, \dots, a_m, \dots, a_M\}, a_m \in \alpha_2 \quad (18)$$

$$\alpha_2 = \left\{ P_{\min}, P_{\min} + \frac{P_{\max} - P_{\min}}{N_p - 1}, \dots, P_{\min} + \frac{(N_p - 2)(P_{\max} - P_{\min})}{N_p - 1}, P_{\max} \right\} \quad (19)$$

其中, a_m 表示基站 m 当前的功率。 α_1 表明基站的决策一被设定为“开启”和“关断”两个等级, α_2 表明基站的决策二被设定为允许开启 N_p 个功率等级。

3) 奖励函数

系统总体优化目标 G 是一个与用户传输时延 d^T 、用户传输能耗 c^T 、基站切换损耗 c^S 和基站恒定能耗

c^{Con} 有关的参数, 它综合考虑了系统能耗和用户通信服务质量。因此, 将奖励函数 R 设计为与 G 负相关的函数, 奖励值取最大值说明当前 G 最小, 即达成系统总体目标最优解 $\min G$ 。奖励函数 R 表示如下:

$$R = -\sum_{m=1}^M \sum_{n=1}^N (\gamma_1 d_{m,n}^T + \gamma_2 c_{m,n}^T + \gamma_3 c_{m,n}^S) - \gamma_4 c^{Con} \quad (20)$$

DDPG 算法在处理大动作空间和连续动作输出时具有一定的优势, 能够采用经验回放、深度神经网络和目标网络等机制, 增强训练过程的稳定性, 加速模型收敛。本文提出的两种基站节能决策中, 动作空间 A 的大小随基站数量 M 呈指数增长, 且动作输出即基站功率等级为确定值。因此, 本文选取 DDPG 作为决策优化方法的基础强化学习算法。

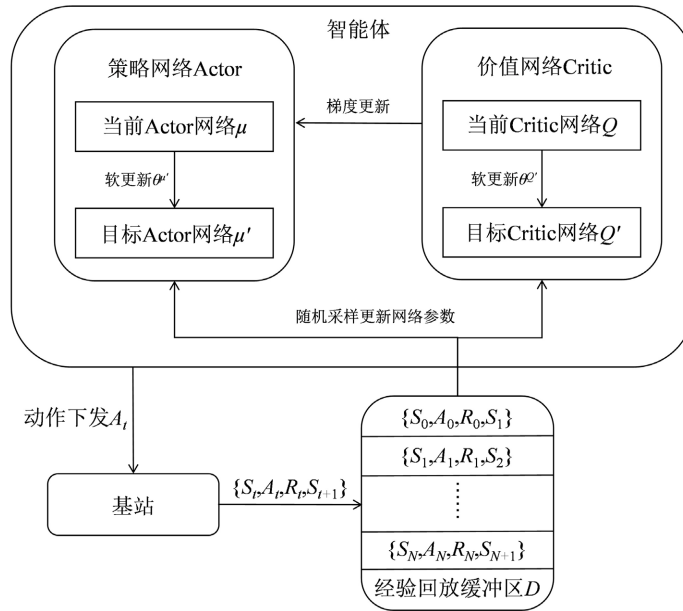


Figure 4. Base station energy-saving decision optimization method based on DDPG

图 4. 基于 DDPG 的基站节能决策优化方法

基于 DDPG 的基站节能决策优化方法如图 4 所示。DDPG 算法由当前 Actor 网络 μ 、当前 Critic 网络 Q 和目标 Actor 网络 μ' 、目标 Critic 网络 Q' 构成。首先, 初始化 Actor 网络参数 θ^μ 和 Critic 网络参数 θ^Q 及其目标网络参数 $\theta^{\mu'}$ 和 $\theta^{Q'}$, 初始化经验回放缓冲区 D 。在每个 step, 各基站采集当前 step 的网络信息, 汇总后形成全局状态空间 S_t , 上报给部署在系统中的智能体。智能体综合全局状态空间 S 输出动作空间 A_t , 下发至各个基站, 基站执行决策。计算该决策下的奖励值 R_t , 获取下一 step 的全局状态空间 S_{t+1} , 将集合 $\{S_t, A_t, R_t, S_{t+1}\}$ 存入经验回放缓冲区 D 。智能体从经验回放缓冲区 D 中随机获取一批样本, 更新网络参数。循环往复直至奖励 R 收敛, 对比节能决策一与决策二下奖励 R 的大小, 从两种决策中选取更优方案。

4. 仿真结果

4.1. 参数设置

仿真实验在面积为 $1 \text{ km} \times 1 \text{ km}$ 的矩形区域范围内进行, 10 个基站均匀分布在其中, 10 个用户在其对应基站附近随机撒点, 分布范围在其半径 100 m 内, 有 4 个用户的通信业务需求随时间降低, 另外 6 个用户保持不变。仿真实验经过 1000 个 episode 的强化学习训练, 每个 episode 包含 50 个 step。设 step

= 0 时每个用户的初始通信业务需求大小为 2000 kbit 到 4000 kbit 不等。具体的仿真平台参数设置如表 1 所示。

Table 1. Simulation parameters

表 1. 仿真参数

参数名称	参数设置
信道模型	Dense Urban 模型 + 阴影衰落
仿真区域	1 km × 1 km
基站数量 M	10 个
基站部署方式	均匀分布
用户数量 N	10 个
用户分布方式	在对应基站半径 100 m 内随机撒点
基站传输带宽 B_m	20 MHz
基站载波频率 f_c	2000 MHz
用户设备发射功率 $p_{m,n}$	最高 23 dBm
基站恒定能耗 c_m^{Con}	满功率: 45 kJ 关断: 5 kJ
基站功率等级 N_P	5 个功率等级
阴影衰落 φ	3 dB
天线增益 ε	14 dBi

4.2. 结果分析

该决策优化方法的目标是降低传输时延、传输能耗、切换损耗和基站的恒定能耗，因此奖励值 R 的绝对值越大，说明系统的时延与能耗性能越差，故将其定义为负值。图 5 和图 6 显示了仿真结果，图中的横坐标为强化学习 DDPG 算法的训练轮次 episode，纵坐标为综合奖励值 R 。

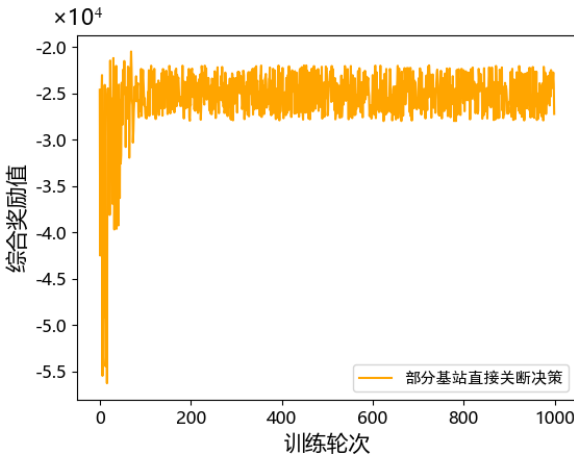


Figure 5. Training performance of DDPG algorithm under energy-saving decision I

图 5. 节能决策一下的 DDPG 算法训练性能

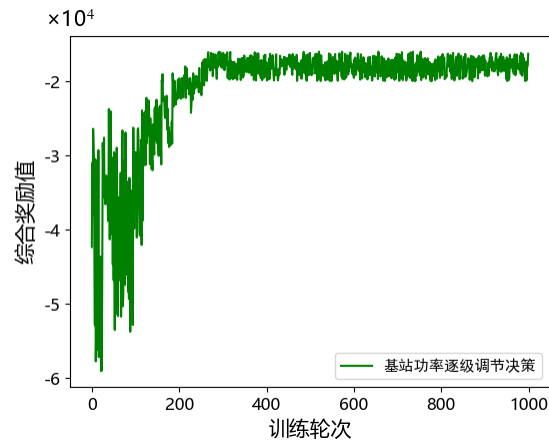


Figure 6. Training performance of DDPG algorithm under energy-saving decision II

图 6. 节能决策二下的 DDPG 算法训练性能

当基站执行节能决策一时，基站仅有“开启”和“关闭”两个功率等级。图 5 表明，训练曲线的收敛速度较快，在第 75 个轮次收敛。这得益于基站执行直接关断动作，用户切换至就近基站，不需要执行细粒度的功率等级调节，基站的恒定能耗迅速减少。当训练曲线收敛后，奖励值稳定在 -2.5×10^4 上下。但基站在执行直接关断决策的情况下，奖励值的波动范围较大，其波动范围在 6×10^3 。这是因为在用户进行基站切换的过程中，附近的基站面对突发新增的通信业务，将出现一些高能耗或高时延的明显抖动。

当基站执行节能决策二时，即基站的功率等级分为 5 档，分别是 5 kJ、15 kJ、25 kJ、35 kJ 和 45 kJ，且能够随着用户通信业务需求的减少逐级降低。图 6 表明，训练曲线的收敛速度相较于执行节能决策一时降低，直到第 260 个轮次才趋于收敛。这是由于细粒度的功率等级要求基站做出更精细化的调节。而当训练曲线收敛后，奖励值稳定在 -2.0×10^4 上下，说明基站逐级调节的节能决策拥有更优的能耗与时延综合性能。其收敛效果相较于节能决策一更稳定，波动范围在 4×10^3 ，这是因为用户无需切换基站，始终保持与初始基站的连接，传输相对稳定。

应用训练完成的智能体模型对部分基站直接关断的节能决策一和基站功率逐级调节的节能决策二进行仿真测试，统计模型收敛后用户在一个 step 下的平均时延与能耗数据如表 2 所示。

Table 2. Comparison of average delay and energy consumption of two energy-saving decisions

表 2. 两种节能决策平均时延与能耗对比

参数指标	节能决策一	节能决策二
平均传输时延 d^T/ms	6.23	4.77
平均传输能耗 c^T/J	1.62×10^{-2}	9.52×10^{-3}
平均切换损耗 c^S/dB	6.16	0

结合图 5、图 6 以及表 2 中的数据可以看出，当基站面临两种节能决策时，执行基站直接关断的决策的模型训练收敛速度更快，即决策响应速度更快，而执行功率逐级调节的决策在降低传输时延、减少传输能耗方面则有着较大优势。

将两种基站节能决策和传统基线方案(即所有基站始终保持满功率运行)的实验结果进行了对比，来验证所提节能决策的有效性。如图 7 所示，两种节能决策的平均奖励值的绝对值均明显低于基线方案。部分基站直接关断的节能决策一相较传统方案性能提升了 47.5%，基站功率逐级调节的节能决策二相较

传统方案性能提升了 56.4%，将所提的基站节能决策相对比，节能决策二的性能比节能决策一高 16.9%。

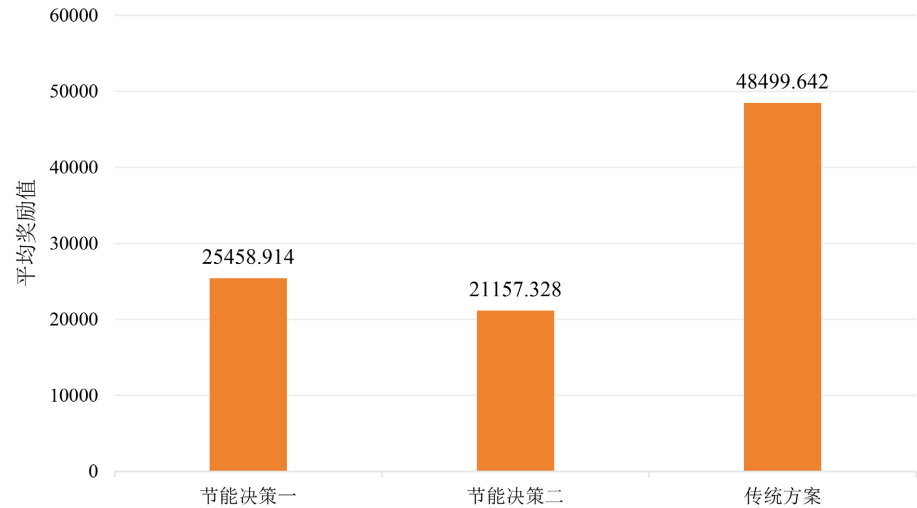


Figure 7. Performance comparison of two energy-saving decisions and baseline
图 7. 两种节能决策和基线之间的性能比较

本文还进一步分析了在基站数量和分布保持不变的情况下，增加用户数量，两种基站节能决策和传统基线方案的平均传输时延以及基站平均能耗的对比，对比结果见图 8 和图 9。

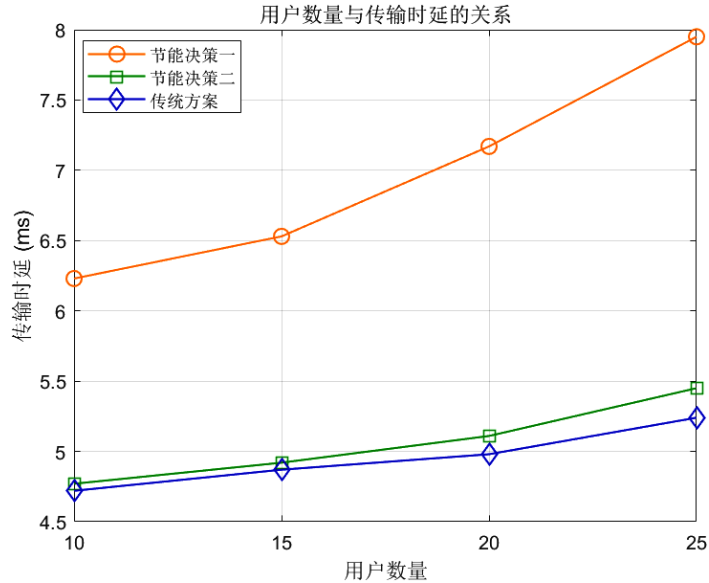


Figure 8. Comparison of average transmission delay under different numbers of users
图 8. 不同用户数量下的平均传输时延对比

由图 8 可知，执行基站直接关断的决策一时，平均传输时延明显高于决策二与传统基线方案。这是由于关断部分基站使得用户切换至剩余基站，仍处于开启状态的基站需要负责更多用户的通信业务，因此传输时延上涨。而随着用户数量的增长，决策一与另外两种决策的时延差距越拉越大，不过差距始终保持在 5 ms 以内，因此仍属于用户可接受的延迟范围。对比功率逐级调节的节能决策二和基站保持满功率运行的传统基线方案，它们的时延表现都相对优异。且在用户数量较少时，决策二和传统方案之间的

时延差距几乎可以忽略不计。这是因为网络场景中用户密度较为稀疏，用户请求的业务量较少，调低基站的功率等级对业务传输质量的影响微乎其微。

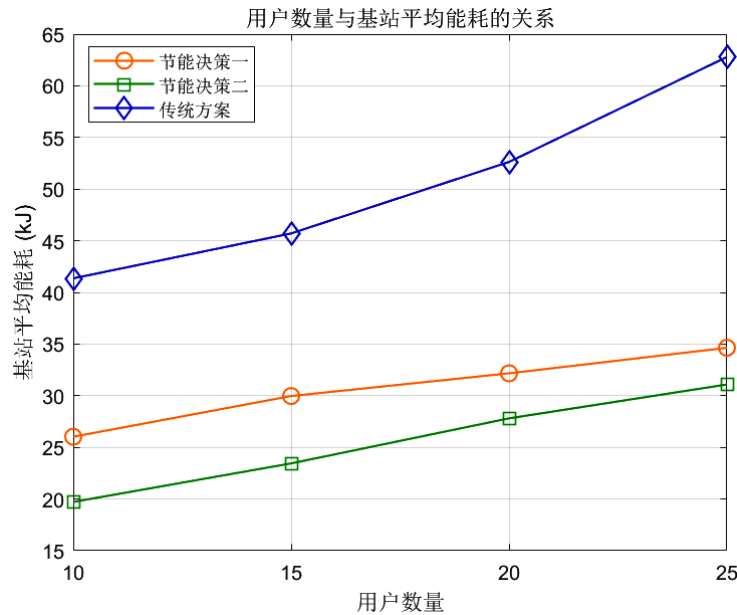


Figure 9. Comparison of average energy consumption of base stations under different numbers of users

图 9. 不同用户数量下的基站平均能耗对比

从图 9 可以看出，当基站执行节能决策一或决策二时，基站的平均能耗始终显著低于传统方案，证明了本文所提出的两种节能决策在降低基站能耗方面的优越性。将两种基站节能决策进行比较可以发现，执行功率逐级降低的决策二时相比直接关断部分基站的决策一，基站的平均能耗更低，这也印证了决策二在时延和能耗整体性能方面的优势。

5. 结论

本文针对用户通信需求随时间降低的研究场景建立数学模型，设计了综合考虑系统能耗和用户通信服务质量的奖励函数，基于 DDPG 算法求解基站节能决策优化问题。仿真结果表明，本文所提出的两种决策的节能效果相较于基站保持满功率运行的传统方案都有显著提升，为降低系统能耗和时延，基站选择优先执行功率逐级调节的节能决策。

参考文献

- [1] 李丽锦. 基于用户关联的基站休眠策略研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2022.
- [2] 费玲玲. 5G 通信基站供电节能技术应用及智能化分析[J]. 信息与电脑(理论版), 2024, 36(22): 140-142.
- [3] 于昌平, 王琦, 吴舒坦, 等. 考虑负荷时空迁移的 5G 基站与配电网协同优化运行[J]. 电力自动化设备, 2024, 44(12): 195-203.
- [4] Peng, J., Hong, P. and Xue, K. (2014) Stochastic Analysis of Optimal Base Station Energy Saving in Cellular Networks with Sleep Mode. *IEEE Communications Letters*, **18**, 612-615. <https://doi.org/10.1109/lcomm.2014.030114.140241>
- [5] Chang, P. and Miao, G. (2016) Joint Optimization of Base Station Deep-Sleep and DTX Micro-Sleep. 2016 *IEEE Globecom Workshops (GC Wkshps)*, Washington, 4-8 December 2016, 1-6. <https://doi.org/10.1109/glocomw.2016.7848943>
- [6] Vereecken, W., Deruyck, M., Colle, D., Joseph, W., Pickavet, M., Martens, L., *et al.* (2012) Evaluation of the Potential

-
- for Energy Saving in Macrocell and Femtocell Networks Using a Heuristic Introducing Sleep Modes in Base Stations. *EURASIP Journal on Wireless Communications and Networking*, **2012**, 1-14.
<https://doi.org/10.1186/1687-1499-2012-170>
- [7] Cai, S., Che, Y., Duan, L., Wang, J., Zhou, S. and Zhang, R. (2016) Green 5G Heterogeneous Networks through Dynamic Small-Cell Operation. *IEEE Journal on Selected Areas in Communications*, **34**, 1103-1115.
<https://doi.org/10.1109/jsac.2016.2520217>
 - [8] Song, Y., Lim, S.H. and Jeon, S. (2023) Handover Decision Making for Dense HetNets: A Reinforcement Learning Approach. *IEEE Access*, **11**, 24737-24751. <https://doi.org/10.1109/access.2023.3254557>
 - [9] Li, Z., Wang, Y., Fan, Y., Wang, Z., Lin, S. and Sun, L. (2024) Base Station Energy Saving Based on Imitation Learning in 5G Network. *Journal of Physics: Conference Series*, **2829**, Article 012012.
<https://doi.org/10.1088/1742-6596/2829/1/012012>
 - [10] Barbarossa, S., Sardellitti, S. and Di Lorenzo, P. (2014) Communicating While Computing: Distributed Mobile Cloud Computing over 5G Heterogeneous Networks. *IEEE Signal Processing Magazine*, **31**, 45-55.
<https://doi.org/10.1109/msp.2014.2334709>