arXiv:2602.04808v1 [cs.IT] 4 Feb 2026

联合休眠模式激活与动态小区负载下的负载均衡：一种组合多臂老虎机方法

Wajahat Bashir Gilkar 和 Gourab Ghatak，*IEEE*会员电气工程系，印度理工学院德里分校，新德里-110016，印度 电子邮箱：eez238361@iitd.ac.in, gghatak@ee.iitd.ac.in

摘要—**我们提出一种组合多臂老虎机建模方法，以机会性地触发 gNode-B（gNB）小小区的休眠模式，随后执行基于小区范围扩展（CRE）的负载均衡流程。其实现方式是确保用户设备的第五代（第五代移动通信）服务质量标识符（5QI）要求得以维持。关键挑战在于，关闭给定的小型小区 gNB 虽能降低其自身能耗，但可能增加相邻 gNB 以及宏 gNB（覆盖小区）上的负载，从而影响整体能效。通过建模动态小区负载，可准确刻画这一现象；该模型综合考虑用户设备的位置、其相对于所有小小区的相对位置以及其数据需求。我们通过实验表明，所提出的组合上置信界（CUCB）随后配合负载均衡器，不仅优于任意使所有小小区保持开启等朴素策略，也优于其他最先进的强化学习方案。所提出的算法可作为开放式无线接入网（O-RAN）近实时（NRT）RAN 智能控制器（RIC）x Apps 来实现。**

索引术语—**高级休眠模式、异构网络、能效（EE）、小小区、业务卸载和第五代移动通信。**

I. 引言

在3GPP标准化的第五代移动通信异构网络（HetNets）中超密集部署新空口（NR）小小区（SCs），是移动网络运营商为降低路径损耗、提高小区边缘用户设备（UE）吞吐量而普遍采用的策略，同时满足ITU-IMT 2020对容量和能效的要求 [1]。然而，随着小小区密度的增加，网络组件的数量也随之增加，从而导致更高的功耗（PC）和排放。自然，移动网络运营商（MNOs）关注高功耗问题，并正在研究以能效（EE）为核心、通过新的网络规划和优化方法加以解决的方案 [2]。需要指出的是，基站（BSs）消耗了整个网络能量的60%至80% [3], ，使其成为无线接入网（RAN）中最耗电的网元。优化其功耗可能会显著降低网络的功耗并提升其能效。

在现有文献中，已有多种降低功耗（PC）的技术，其中小小区（SC）休眠模式策略尤为吸引人，因为它无需改变网络架构 [4]。要高效实施小蜂窝休眠策略，必须考虑时空网络负载，因为用户设备（UE）的服务质量（QoS）与网络的节能之间存在固有的权衡。在此背景下，我们考虑一种控制/数据分离架构（CDSA） [5], ，其中宏小区（MC）充当控制基站，以确保连续覆盖，同时处理信令功能和低数据速率服务。另一方面，小小区（SC）作为容量小区提供高数据速率服务，并通过适当的回传连接到宏小区。随后，我们提出一种组合多臂老虎机（CMAB）模型，用于形式化 gNB小小区的小区去激活/激活决策（与 3GPP TS 28.310 中的节能管理功能保持一致），并结合负载均衡过程（通过小区范围扩展（CRE）优化），这些过程可通过开放式 RAN 近实时（NRT）RAN 智能控制器（RIC）xApps 实现。作为解决方案，我们采用组合上置信界（CUCB）算法来确定小小区的活动状态。

相关工作： [6]的作者研究了密集多层网络中的用户设备关联与小小区休眠问题，并将其建模为复杂的混合整数规划问题。他们提出了两种低复杂度启发式算法，用于确定最优的用户设备关联和小小区切换模式。在 [7], 中，采用贪心启发式算法来寻找最优的小小区开关配置。启发式算法通常计算开销较大，因此适用范围有限。相比之下，强化学习（RL）方法在解决这一问题上日益受到关注。例如，[8], 利用Q学习框架，在双层异构网络中确定了最优的切换和流量卸载策略。类似地，Salem等人在 [9] 中采用Q学习算法，确定小小区应在特定休眠级别停留的时长，从而在节能与时延之间实现公平权衡。Ozturk 等人 [10]采用SARSA与价值函数近似研究了功耗最小化问题，同时保证用户设备的服务质量。从多臂老虎机（MAB）的角度出发， [11] 研究了联合节能与干扰协调问题，并提出了一种两级算法，用于确定能效最高的控制动作。将经典MAB技术用于这一问题的主要挑战在于动作集具有组合特性，且其规模会随着部署密度的增加而急剧增长。

研究空白与贡献： 上述研究均未考虑小小区和宏小区的动态小区负载来建模功耗。具体而言， [12] 中的作者指出，小区的时空负载不仅取决于

连接的用户设备数量及其相应的数据需求，还取决于这些用户设备相对于服务基站和干扰基站的位置。因此，关闭某个小小区对整体网络负载的影响，取决于连接到该小小区的用户设备相对于其他小小区以及可能接收其卸载流量的宏小区的位置。我们是首个将此类动态小区负载纳入考量以优化休眠模式决策的研究。总体而言，与以往文献相比，本文的贡献如下：

- 小小区休眠策略被建模为一个组合多臂老虎机问题，其超臂奖励取决于动态小区负载，从而导致非单调奖励函数——这是现有文献尚未解决的问题。针对这一问题，本文提出了一种针对单调奖励进行优化的CMAB算法，以寻找最佳休眠配置。结果证实，即使在这种不存在可证明解的非单调场景下，该算法仍然有效。
- 获得最优ASM配置后，可以通过与gNB小小区E2节点连接的O-RAN近实时无线智能控制器xApps加以实施，并在多个流量间隔内保持有效。为进一步提升能效，本文提出了一个负载均衡模块，该模块可以映射到3GPP TS 32.500 [13]中定义的自组织网络（SON）功能。具体而言，通过优化保持开启的小小区的小区范围扩展参数来实现这一目标。
- 我们将本文方案与 [10]提出的ALL-ON策略和基于VFA-SARSA的策略进行了比较，结果表明，本文提出的方案优于这两种策略。出现这一结果的原因在于考虑了动态小区负载和CRE优化。

II. 动态小区负载与功耗

网络几何与传播： 考虑一个超密集 RAN，其中在一个 MC 的覆盖区域内部署了 L个同频 SC。在第五代移动通信及后续架构中，MC 作为覆盖小区提供机会式数据业务，而 SC 则作为容量小区。O-RAN 近实时无线智能控制器通过 E2接口动态配置 SC gNB 的 CRE 偏置，以优化用户设备关联并执行负载均衡，符合 3GPP 自组织网络（SON）和 TS 32.500 [13] 中的无线资源管理（RRM）功能。设 ϕ_i 为 $i-th$ 个 SC 的 CRE 参数，ϕ_M 为 MC 的 CRE 参数。设 SC 和 MC 的发射功率分别为 P_S 和 P_M。此外，考虑路径损耗指数为 α_p 、SC 和 MC 的路径损耗常数分别为 K_S 和 K_M 的 3GPP 路径损耗模型。再设用户设备与 SC i之间的相对距离为 d_{ui}，而用户设备与 MC 之间的距离为 d_{uM}。用户设备关联基于参考信号接收功率（RSRP），因此用户设备 u 在以下情况下连接到第 i 个 SC：

$$\phi_i P_S K_S d_{ui}^{-\alpha_p} > \max_{l\neq i}\{\phi_l P_S K_S d_{ul}^{-\alpha_p}, \phi_M P_M K_M d_{uM}^{-\alpha_p}\}.$$

如果用户设备 u 从 MC 接收到的接收功率最高，即，则该用户设备与 MC 关联。

$$\phi_M P_M K_M d_{uM}^{-\alpha_p} > \phi_i P_S K_S d_{ui}^{-\alpha_p},\ \forall\, i \in \{1,2,....,L\}.$$

连接到宏小区的第 j 个用户设备所经历的信干噪比 (SINR) 为

$$\chi_{jM} = \frac{P_M K_M d_{jM}^{-\alpha}}{N_0 + \sum_{l=1}^{L} P_S K_S d_{jMl}^{-\alpha}\delta_l} \tag{1}$$

其中，分母中的第二项表示该用户设备受到网络中小小区的干扰。d_{jMl} 是连接到宏小区的第 j 个用户设备与第 l 个小小区之间的距离，N_0 表示噪声功率，$\delta_l \in \{0,1\}$表示第 $l-th$ 个小小区的状态，即开启还是关闭。类似地，与第 i 个小小区关联的第 j 个用户设备所经历的信干噪比为

$$\chi_{ji} = \frac{P_S K_S d_{j,i}^{-\alpha}}{N_0 + P_M K_M d_{jiM}^{-\alpha} + \sum_{l=1,l\neq i}^{L} P_S K_S djil^{-\alpha}\delta_l} \tag{2}$$

其中，分母中的第二项和第三项分别表示该用户设备受到网络中宏小区和其他小小区的干扰。

业务模型： 设网络中共有 N 个用户设备，其中 $N = N_M + \sum_{i=1}^{L} N_i$。此处，$N_M$ 和 N_i 分别为由宏小区和第 i 个小小区服务的用户设备数量。此外，用 B_M 和 B_S 分别表示宏小区和每个小小区的带宽。用户请求按平均到达率为 λ_a /s/m^2 的泊松到达过程到达。每个用户请求的数据大小为 σ 位，因此业务量为 $\omega = \sigma\lambda_a$ 位/s/m^2。连接到宏小区的用户设备所经历的下行数据速率因此为 $R_{jM} = \frac{B_M}{N_M}\log(1+\chi_{jM})$。因此，该用户设备在宏小区产生的负载为

$$\rho_{jM} = \frac{\omega}{R_{jM}} = \frac{\omega N_M}{B_M \cdot \log(1+\chi_{jM})} \tag{3}$$

因此，所有与宏小区关联的 N_M 用户设备所产生的总负载为

$$\rho_M = \sum_{j=1}^{N_M} \rho_{jM} = \sum_{j=1}^{N_M} \frac{\omega \cdot N_M}{B_M \cdot \log(1+\chi_{jM})}. \tag{4}$$

注 1. 需要强调的是，这种小区负载的表述方式不同于其他研究异构网络能量管理的工作。具体而言，这种表述不仅考虑了已连接用户设备的数量及其数据请求，更重要的是，还考虑了它们的位置和所经历的信干噪比。这更准确地描述了实际的动态小区负载。

同样，与第 i 个小区关联的第 j 个用户设备的数据速率为 $R_{ji} = \frac{B_S}{N_i}\cdot\log(1+\chi_{ji})$，因此产生相应的小区负载 $\rho_{j,i} = \frac{\omega}{R_j} = \frac{\omega N_i}{B_S\cdot\log(1+\chi_{ji})}$。因此，第 i 个小型小区上的总负载为

$$\rho_i = \sum_{j=1}^{N_i} \rho_{ji} = \sum_{j=1}^{N_i} \frac{\omega \cdot N_i}{B_S \cdot \log(1+\chi_{ji})}. \tag{5}$$

功耗模型根据能效无线电与网络技术（EARTH）功耗模型 [14] , ， $i-th$ 个小型小区的瞬时功耗 ξ_i 为

$$\xi_i = \begin{cases} P_{oi} + \eta_i \rho_i P_S, & \delta_i = 1, \\ P_{\text{sleep}}, & \delta_i = 0 \end{cases} \tag{6}$$

其中， P_{oi} 和 P_{sleep} 分别是第 i个小小区的运行电路功耗和休眠电路功耗， η_i 是功率放大器（PA）的效率。由于宏小区是覆盖小区，因此始终处于开启状态。宏小区的功耗为 $\xi_M = P_{oM} + \eta_M \rho_M P_M$，其中 P_{oM}、 η_M、 ρ_M和 P_M分别是宏小区的运行电路功耗、PA 效率、负载因子和发射功率。网络的瞬时功耗 ξ_T 为

$$\xi_T = \xi_M + \sum_{i=1}^{L} \xi_i. \tag{7}$$

III. 问题构建与 CMAB 框架

目标是找到一种休眠配置 $\{\delta_1, \delta_2, \ldots, \delta_L\} \in \{0,1\}^L$，使网络功耗最小，同时确保网络中的所有用户设备都能获得有保障的比特率 R_{min}。请注意，每当小小区的状态发生变化（例如从 1 变为 0）时，必须根据次优 RSSI 测量值，将连接到该小小区的用户设备预先卸载到网络中其他处于开启状态的小小区或宏小区。这不仅会改变新小区的负载，还会重新配置整个网络的能耗。此外，这还会改变转移后用户设备所体验到的信干噪比。小小区的配置从 0 变为 1 时也会发生同样的情况。我们将目标优化问题表述为

$$\begin{aligned} \min_{\{0,1\}^L} \quad & \xi_T \\ \text{s.t.} \quad & \rho_M \le 1; \ \ \rho_i \le 1, \ \ \forall i \in \{1,2,....,L\} \\ & R_{jM} \ge R_{\min}; \ \ R_{ji} \ge R_{\min}, \ \ \forall i \in \{1,2,....,L\} \end{aligned}$$

第一个约束确保网络中的所有小区都不会过载。最后一个约束确保向用户设备提供最低保证比特率。

为了解决上述问题，我们在 CMAB 框架中重构该问题，因为经典的在线学习或老虎机模型无法随着小小区数量的增加而扩展。为此，我们将每个小小区的状态视为一个臂（此处为简单臂）。在 CMAB 中，玩家选择的不是一个简单臂，而是一组称为超级臂的简单臂，该超级臂会为玩家带来超级臂奖励。在经典的 CMAB 问题中，超级臂奖励取决于其组成简单臂的奖励。选择一个超级臂后，该超级臂所包含的简单臂的奖励也会随之揭示，这些奖励将用于后续轮次的超级臂选择。因此，目标是在一段时间内最大化累积超级臂奖励。

备注 2。 与经典的 *CMAB* 设置不同，在我们的问题中，选择超级臂所获得的奖励不仅取决于其组成的简单臂，还取决于因对其他臂造成不利影响而产生的惩罚项。这进一步意味着现有 *CMAB* 算法（例如，参见 *[15]*）中基于单调性假设得出的遗憾保证可能并不成立。

算法 1 α， β-近似预言机

1: 输入： 简单臂
臂奖励的均值向量， $\alpha \in [0,1]$
2: 对于每个超级臂 $S \in \mathcal{S}$ **执行**
3: 使用向量 $\bar{\mu}$计算$r_{\bar{\mu}}(S)$。
4: **结束循环**
5: 计算相对于
给定的均值向量 $\bar{\mu}$:
6: $r_{\bar{\mu}}(S^*) = $ 最大$_{S \in \mathcal{S}} r_{\bar{\mu}}(S)$
7: 确定 α–最优超级臂集合，如下:
8: $\mathcal{S}' = \{S \in \mathcal{S} : r_{\bar{\mu}}(S) \ge \alpha . r_{\bar{\mu}}(S^*)\}$
9: 以概率 β 返回超级臂 $S \in \mathcal{S}'$。

我们对 CMAB 问题进行数学表述。令简单臂对应于为一个小型小区 l设置 $\delta_l = 0$ 。此外，玩家可以选择在某个时间步不玩任何臂，即 $\delta_l = 1, \ \forall l \in [L]$。在每个时隙，简单臂 μ_l 的平均奖励定义为

$$\mu_l = \mathbb{E}\left[P_{l,\text{all}} - P_{\text{sleep}}\right], \tag{8}$$

其中， $P_{i,\text{all}}$ 表示所有 SC 均开启时第 l 个 SC 的功耗。期望是针对 ω 中的随机性而言的。

备注 3. 上述平均奖励的表述，从静态决策的角度比较了一个简单动作的性能——即让所有 *SC* 保持开启。这不同于功率节省方面的瞬时收益，也就是与决策前的功耗相比所节省的功率。此形式化定义使问题环境变得无状态且平稳，从而能够应用 *CMAB* 算法。

执行一个超级臂时，定义 $\Delta P_M = \eta_M P_M(\rho_{M,\text{all}} - \hat{} \rho_M)$ 为 MC 的功率变化，定义 $\Delta P_i = \eta_i P_i(\rho_{i,\text{all}} - \hat{} \rho_i)$ 为 $i-th$ SC 的功率变化，其中 $\hat{}\rho_M$ 和 $\hat{}\rho_i$ 分别是执行超级臂时 MC 和 $i-th$SC 的估计负载值，而 $\rho_{M,\text{all}}$ 和 $\rho_{i,\text{all}}$ 分别是网络实施全开策略时 MC 和 $i-th$ SC 的负载值。基于此，我们将超级臂 $S \subseteq [K]$ 的期望奖励定义为

$$\begin{aligned} r_\mu(S) = & \underbrace{\sum_{i=1}^{L} \mu_i (1-\delta_i)}_{A} + \underbrace{\Delta P_M}_{B} + \underbrace{\sum_{i=1}^{L} \Delta P_i \delta_i}_{C} \\ & \underbrace{- \mathcal{P}_1 \max(0, (\hat{\rho}_M - 1)) - \sum_{i=1}^{L} \mathcal{P}_2 \max(0, (\hat{\rho}_i - 1))}_{D} \end{aligned} \tag{9}$$

项 A 表示被关闭的小小区所带来的功率节省。项 B 是因关闭部分小小区而导致的宏小区功耗增减。项 C 表示保持开启的小小区功耗的增减。最后，项 D 表示对过载判据的违反，其中 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 为惩罚项。

A. (α, β)近似预言机与 *CUCB*

回想一下，由于该组合问题具有NP难性质，大多数 CMAB 算法在动作选择时依赖一个 α, β近似预言机。具体而言，对于某个 $\alpha, \beta \leq 1$，算法输出一个超级臂 S，使得$\mathbb{P}(r_\mu(S) \geq \alpha . r_\mu(S^*)) \geq \beta$，其中$r_\mu(S^*) = \max^{S \in \mathcal{S}} r_\mu(S)$为算法1所示的最优超级臂奖励。在此表述下，由于仅有 β 比例的预言机计算成功，且成功时奖励仅为最优值的 α近似，性能以所定义的 α, β近似遗憾为基准

$$\mathrm{Reg}_{\mu,\alpha,\beta}(n) = n\alpha\beta r_\mu(S^*) - E_S\left[\sum_{t=1}^{n} r_\mu(S_t)\right]. \qquad (10)$$

现在介绍算法2中所示的 CUCB 算法。经过前 m 轮初始化（轮数等于简单臂的数量，此处即 SC 的数量）后，我们为每个臂维护一个经验均值 $\hat{}\mu_i$。将期望向量 $\hat{}\mu$ 连同附加的置信区间（取决于相应臂被选择的次数， T_i）一起输入到 (α, β)-近似预言机，二者共同构成上置信界。随后，预言机根据均值向量 $\bar{\mu}$返回一个 α-最优超级臂，并选择该超级臂；之后相应地更新变量 T_i's 和 $\hat{}\mu_i$'s。

请注意，由于超臂奖励具有非单调性，(α, β)-近似预言机会穷举评估每个超级臂的奖励。不过，计算这些超臂奖励所需的期望基臂奖励，仍通过利用 CMAB 的结构化框架来学习，从而实现快速学习。

*B. CRE*优化

直接应用 (α, β)-近似预言机返回的超级臂 S 可能并不十分节能

算法2 使用计算预言机的 CUCB

对于每个简单臂 a_i 维护:
（1）变量 T_i 作为臂 a_i 被迄今为止执行的总次数。
（2）变量$\hat{}\mu_i$ 作为所有结果 r_i的估计均值臂 i 迄今为止所观测到的。
对于每个臂 a_i **执行**
　执行一个任意超级臂 $S \in \mathcal{S}$，使得 $i \in S$。
　更新变量 T_i 和 $\hat{}\mu_i$。
结束循环
$t \leftarrow m$。
当 $t \leq n$ **时**
　$t \leftarrow t + 1$。
　对于每个臂 i **执行**
　　集合 $\bar{\mu}_i = \hat{\mu}_i + \sqrt{\frac{3 \ln t}{2T_i}}$。
　结束循环
　$S =$ 预言机$(\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_m)$。
　执行 S 并更新所有 T_i和$\hat{}\mu_i$。
结束循环

算法 3 鲍威尔优化方法

输入： 初始点 $\mathbf{\Phi}_0$，函数 f，容差 ϵ_p，最大迭代次数 N
初始化： 将方向向量 $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_n\}$ 设为单位向量
对于 $k = 1$ 至 N **执行**
　$\mathbf{\Phi}_{start} \leftarrow \mathbf{\Phi}_0$
　对于 $i = 1$ 至 n **执行**
　　执行线搜索：找到使 $f(\mathbf{\Phi}_{i-1} +$ 最小的 $\alpha_i \alpha_i \mathbf{d}_i)$
　　更新： $\mathbf{\Phi}_i \leftarrow \mathbf{\Phi}_{i-1} + \alpha_i \mathbf{d}_i$
　结束循环
　$\mathbf{d}_{n+1} \leftarrow \mathbf{\Phi}_n - \mathbf{\Phi}_0$
　沿 $\mathbf{d}_{n+1}$ 执行线搜索：找到使其最小的 α $f(\mathbf{\Phi}_n + \alpha \mathbf{d}_{n+1})$
　更新： $\mathbf{\Phi}_{new} \leftarrow \mathbf{\Phi}_n + \alpha \mathbf{d}_{n+1}$
　如果 $\|\mathbf{\Phi}_{new} - \mathbf{\Phi}_0\| < \epsilon_p$ **则**
　　中断
　结束如果
　更新方向集：将一个$\mathbf{d}_i$ 替换为$\mathbf{d}_{n+1}$
　$\mathbf{\Phi}_0 \leftarrow \mathbf{\Phi}_{new}$
结束循环
输出： 近似最小点 $\mathbf{x}_{new}$

因为负载均衡效率低下，可能会增加宏小区以及保持开启的小小区的负载。因此，为了进一步节能，我们提出根据超级臂 S 优化保持开启的小小区的 CRE 参数。CRE 优化问题表述为

$$\min_{\Phi} \rho_M + \sum_{i \in S} \rho_i \delta_i$$
$$\text{s.t} \quad \rho_M \leq 1;\ \rho_i \leq 1, \forall\ i \in S.$$

首先设置 $\phi_i = 0,\ \forall i \in S$。然后注意到，计算目标函数相对于 Φ 的导数不可行。为了解决这一问题，我们采用无导数的 Powell 共轭方向法（如算法 3 所示），在不使用函数 $f: \mathbb{R}^n \to \mathbb{R}$的导数 [16]的情况下寻找其局部最小值。该算法沿一组 n初始时相互正交的方向，依次进行一维最小化（使用线搜索）。算法以初始点 Φ_0 和一组搜索方向向量 $\{d_1, d_2, ..., d_n\}$初始化。从 Φ_0开始，算法沿每个方向 d_i依次最小化 f，并在每次线搜索后更新当前点。遍历完所有 n个方向后，定义一个等于起点净位移的新方向，并沿该方向额外执行一次线搜索；如果此次最小化的变化非常小（小于 ϵ_p），算法便停止。否则，算法用新的位移方向替换一个旧方向，并重复这一过程，最多执行 N次迭代，从而自适应地构造近似共轭方向并高效收敛。

IV. 数值结果

我们的仿真框架由一个 MC 覆盖圆盘构成，其中 SC 和用户设备均匀部署。网络

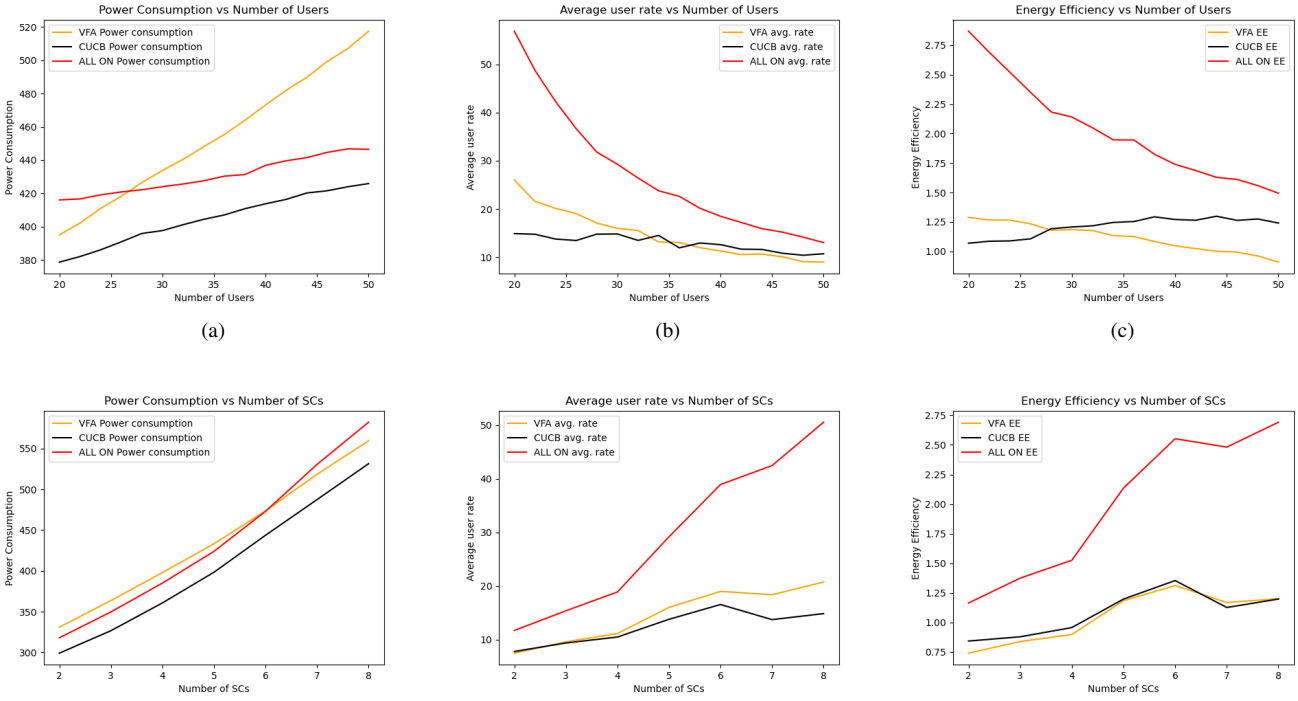


图 1. (a) 5 个小小区下功耗随用户设备数量的变化，(b) 用户设备平均数据速率，(c) 能效。(d) 30 个用户设备下功耗随小小区数量的变化，(e) 用户设备平均数据速率，(f) 能效。

表 I：仿真参数

参数	数值
宏小区半径	500 米
宏小区/小小区载波频率	2.0 GHz
宏蜂窝/小蜂窝带宽	20 MHz
路径损耗指数（α_p）	4
噪声功率谱密度	3.98 x 10^{-21} 瓦/赫兹
$\mathcal{P}_1$, $\mathcal{P}_2$	100（两者）
(α, β)-近似预言机	α = 0.989, β = 0.98

仿真参数见表 I，表 II 则列出了各种类型基站的功耗特性。仿真中我们仅使用微型小蜂窝，但其他类型的小蜂窝也会得到类似结果。我们将所提出机制的性能与全开策略（不实施小蜂窝切换）以及 [10]，中提出的小蜂窝休眠策略进行比较，后者称为 VFA-SARSA。VFA-SARSA 方法利用 CDSA 架构，并将小蜂窝切换问题建模为一个受服务质量约束的能量最小化问题。由于状态–动作空间会随小蜂窝数量呈指数级增长，因此采用一种称为 SARSA（状态–动作–奖励–状态–动作）的强化学习算法，并结合价值函数近似（VFA），以寻找最优开/关模式。此外，我们假设这两种方法均未进行 CRE 参数优化。图 1a 显示，随着用户设备数量增加，基站负载随之上升，各策略的功耗也相应增加。用户设备数量不超过 27 时，VFA-SARSA 的表现优于全开策略，

表 II：基站的功耗曲线 [14]

基站类型	效率 η	功耗 [W]		
		P_S	P_{oi}	P_{sleep}
宏	4.7	20	130	75
RRH	2.8	20	84	56
微	2.6	6.3	56	39
Pico	4.0	0.13	6.8	4.3
飞	8.0	0.05	4.8	2.9

随后，由于缺乏适当的负载均衡，它开始比ALL-ON消耗更多功率。事实上，负载相关功耗的增加并未被关闭部分小小区所带来的功率节省所抵消。CUCB通过在小型小区开启/关闭决策之后采用高效的负载均衡优化，表现优于其他方法。随着用户设备数量增加，CUCB策略的表现接近ALL-ON，但仍消耗更少功率。图1b表明，由于没有进行小小区开关切换，ALL-ON最初提供的数据速率高于另外两种策略。在用户设备数量较少时，VFA-SARSA提供的数据速率也优于CUCB。但一旦用户设备数量增加，CUCB便优于VFA-SARSA，因为它能够通过高效优化网络负载来管理更多用户设备。随着用户设备数量增加，CUCB和VFA-SARSA都开始趋近于ALL-ON。由于在用户设备密集时，CUCB在数据速率和功耗两方面均优于VFA-SARSA，其能效随用户设备数量增加而提高，如图1c所示。当网络中小小区数量更多时，也观察到类似的性能趋势。由上述分析可以得出，

CUCB能够比VFA-SARSA和ALL-ON管理更多的用户设备，从而在不消耗更多功率的情况下提高网络容量。

图1d显示，对于这三种策略，随着部署的小型小区数量增加，网络功耗都会上升。CUCB的功耗低于其他策略；而VFA-SARSA起初（小型小区数量较少时）由于负载较高，功耗高于ALL-ON。在宏小区覆盖区域内，小型小区数量达到6个后，这一情况得到改善。如图1e所示，随着网络致密化程度提高，节能效果提升的代价是数据速率降低。尽管如此，仍满足1 MBps的服务质量要求。CUCB通过关闭负载较低的小型小区来降低功耗，这会增加其他开启的小型小区和宏小区的负载，从而导致用户设备的数据速率降低。VFA-SARSA也有同样的情况，但其用户设备平均数据速率略高。ALL-ON的用户设备平均数据速率优于其他所有策略，因为所有小型小区都处于开启状态且负载较低，这表明此处干扰的影响小于增加基站数量带来的好处。图1f展示了能效随小型小区数量变化的趋势。ALL-ON的表现优于其他策略，因为它能提供更高的用户设备平均数据速率。小型小区数量较少时，CUCB策略的能效优于VFA-SARSA，因为前者提供的用户设备平均数据速率几乎相同，但功耗更低，如图1d和图1e所示。随后，随着小型小区数量增加，CUCB和VFA-SARSA的能效趋于一致。当网络中的用户设备数量较多时，无论在数据速率还是节能方面，CUCB都优于VFA-SARSA。

未来工作将研究非均匀用户设备流量分布下的性能。

V. 系统设计见解

- **集群激活策略：**结果表明，最优的小区休眠模式与时空业务量模式相关。将负载曲线相关的小小区分组并应用 CUCB 策略，可以显著降低切换开销，同时保持性能。
- **探索预算选择：**观察到的次线性遗憾行为表明，有限的探索阶段（最好安排在低流量时段）足以实现收敛，从而以最小的性能损失实现稳定的长期运行。
- **CRE 与休眠联合优化：**基于 Powell 方法的 CRE 调优显著提升了小区休眠期间宏小区的卸载效率。将 CRE 调整与小区激活相协调，与静态 CRE 设置相比，能够实现更好的负载均衡并持续保障用户服务质量。
- **对流量变化的鲁棒性：**该方法能够有效适应随时间变化的业务量，这表明无需依赖精确的流量预测即可部署该方法，只需使用在线测量进行策略更新。

VI. 结论

本文提出了一种赌博机框架，用于切换小小区的开启/关闭状态，并高效优化超高密度异构网络的负载，以降低低业务量时段的功耗。该机制采用一种称为 CUCB 的 CMAB 算法，学习新空口小区的最佳开启/关闭配置。决策输出可集成到 RIC 中的节能管理功能。随后，执行一种无导数连续优化过程，根据 3GPP 无线资源管理规范调节 CRE 参数，从而优化网络中的整体负载分布。这不仅节省了能量，还提升了网络的负载处理能力。我们将该机制与全开策略以及一种称为 VFA-SARSA 的先进技术进行了比较。仿真结果表明，在保证用户设备所需服务质量的同时，所提出的 CMAB 解决方案相比全开策略和 VFA-SARSA 策略消耗更少的功率。此外，结果还表明，与 VFA-SARSA 相比，所提出的方法能够处理数量更多的用户设备，从而提升网络容量。

参考文献

[1] F. Kooshki等人，“无小区无线接入网中面向第五代移动通信及超第五代移动通信网络的高能效休眠模式方案”，IEEE Access，第11卷，第1432–1444页，2022年。[2] E. Kolta等人，“走向绿色：移动通信能效基准测试”，GSMA Intell，2021年。[3] M. Feng等人，“第五代移动通信无线网络中的基站开关：方法与挑战”，IEEE Wirel. Commun.，第24卷，第4期，第46–54页，2017年。[4] J. Wu等人，“绿色蜂窝网络中高能效基站休眠模式技术综述”，IEEE Commun. Surv. Tutor.，第17卷，第2期，第803–826页，2015年。[5] A. Mohamed等人，“蜂窝无线接入网的控制与数据分离架构：综述与展望”，IEEE Commun. Surv. Tutor.，第18卷，第1期，第446–465页，2015年。[6] Y. Sun和Y. Chang，“多层超密集小蜂窝网络中的高能效用户关联与小区休眠策略”，Electronics Letters，第54卷，第12期，第787–789页，2018年。[7] J. Wu等人，“面向绿色超密集网络的小蜂窝动态开关控制”，见第8届无线通信与信号处理国际会议论文集（WCSP），IEEE，2016年，第1–5页。[8] A. I. Abubakar等人，“Q学习辅助的异构网络能量感知业务量卸载与小区切换”，见2019年IEEE第24届通信链路与网络计算机辅助建模与设计国际研讨会（CAMAD），IEEE，2019年，第1–6页。[9] F. E. Salem等人，“第五代移动通信网络中高级休眠模式管理的强化学习方法”，见IEEE第88届车辆技术会议论文集（VTC-Fall），IEEE，2018年，第1–5页。[10] M. Ozturk等人，“通过流量感知小区切换实现超密集无线接入网中的能量优化”，IEEE Trans. Green Commun. Netw.，第5卷，第2期，第832–845页，2021年。[11] J. A. Ayala-Romero等人，“异构网络中用于节能与干扰协调的在线学习”，IEEE J. Sel. Areas Commun.，第37卷，第6期，第1374–1388页，2019年。[12] G. Ghatak等人，“噪声受限随机蜂窝网络中动态小区负载的精确表征”，见2018年IEEE第88届车辆技术会议（VTC-Fall），IEEE，2018年，第1–5页。[13] L. Jorguseski等人，“3GPP中的自组织网络：标准化与未来趋势”，IEEE Commun. Mag.，第52卷，第12期，第28–34页，2014年。[14] G. Auer等人，“运行无线网络需要多少能量？”，IEEE Wirel. Commun.，第18卷，第5期，第40–49页，2011年。[15] W. Chen等人，“组合多臂老虎机：通用框架与应用”，见第35届国际机器学习会议，PMLR，2013年，第151–159页。[16]M. J. Powell，“一种无需计算导数即可求多元函数最小值的高效方法”，The Computer Journal，第7卷，第2期，第155–162页，1964年。